GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:12 ; Search time 38.7752 Seconds
(without alignments)
711.353 Million cell updates/sec

Title: US-09-917-791-21
Perfect score: 1071
Sequence: 1 IKVNNWDLFFSPSEDNFTND.....NIGNMLYKDDFVGALIFSGA 207

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1071	100.0	207	22	AAG79295	Amino acid sequenc
2	1071	100.0	407	21	AAY77141	Native botulinum n
3	1071	100.0	413	22	AAB04168	Botulism toxin hea
4	1071	100.0	413	23	ABG69074	Botulinum neurotox
5	1071	100.0	423	21	AAB36302	C. botulinum BoNT/
6	1071	100.0	837	21	AAY77140	Native botulinum n
7	1071	100.0	861	23	ABG69075	Botulinum neurotox
8	1071	100.0	871	19	AAW56019	Recombinant botuli
9	1071	100.0	871	19	AAW56007	Recombinant botuli
10	1071	100.0	871	19	AAW56008	Botulinum neurotox
11	1071	100.0	873	19	AAW56016	Recombinant botuli
12	1071	100.0	875	19	AAW56009	Recombinant botuli
13	1071	100.0	878	19	AAW56010	Recombinant botuli
14	1071	100.0	894	19	AAW56015	Recombinant botuli
15	1071	100.0	907	19	AAW56012	Recombinant botuli
16	1071	100.0	953	19	AAW56011	Recombinant botuli
17	1071	100.0	1013	19	AAW56013	Recombinant botuli
18	1071	100.0	1067	21	AAY93307	A manganese supero
19	1071	100.0	1092	21	AAY93310	A manganese supero
20	1071	100.0	1295	23	AAU99339	Clostridium botuli
21	1071	100.0	1296	17	AAR95010	C. botulinum type
22	1054.5	98.5	847	22	AAB04081	Botulism toxin hea
23	377	35.2	410	22	AAB04102	Botulism toxin hea
24	368.5	34.4	382	22	AAB04100	Botulism toxin hea
25	368.5	34.4	804	23	ABG69083	Botulinum neurotox
26	366.5	34.2	413	22	AAB04097	Botulism toxin hea
27	366.5	34.2	848	22	AAB04082	Botulism toxin hea
28	366.5	34.2	852	23	ABG69077	Botulinum neurotox
29	366.5	34.2	858	19	AAW56018	Recombinant botuli
30	366.5	34.2	1070	21	AAY93308	A manganese supero
31	366.5	34.2	1095	21	AAY93311	A manganese supero
32	366.5	34.2	1169	19	AAW56017	Recombinant botuli
33	364.5	34.0	848	23	ABG69087	Botulinum neuroto
34	364.5	34.0	1059	21	AAY93309	A manganese supero
35	364.5	34.0	1084	21	AAY93312	A manganese supero
36	364.5	34.0	1291	19	AAW68392	Clostridium botuli
37	316	29.5	408	22	AAB04101	Botulism toxin hea
38	316	29.5	858	23	ABG69085	Botulinum neurotox
39	312	29.1	1315	22	AAB61169	Clostridium tetani
40	290.5	27.1	386	22	AAB04099	Botulism toxin hea
41	290.5	27.1	824	23	ABG69081	Botulinum neurotox
42	263	24.6	399	22	AAB04098	Botulism toxin hea
43	263	24.6	811	23	ABG69079	Botulinum neurotox
44	263	24.6	1291	20	AAY05814	Non-toxic modified
45	116.5	10.9	548	19	AAW56014	Recombinant botuli

ALIGNMENTS

RESULT 1
AAG79295
ID AAG79295 standard; Protein; 207 AA.
XX
AC AAG79295;
XX
DT 03-JAN-2002 (first entry)
XX
DE Amino acid sequence of botulinum neurotoxin (BoNT) fragment.
XX
KW Neurotoxin; BoNT; vaccine; botulism.
XX
OS Clostridium botulinum.
XX
PN US6287566-B1.
XX
PD 11-SEP-2001.
XX
PF 19-MAY-1995; 95US-0446114.
XX
PR 19-MAY-1995; 95US-0446114.
XX
PA (USSA) US SEC OF ARMY.
XX
PI Dertzbaugh MT;
XX
DR WPI; 2001-615462/71.
XX
PT New protective epitopes of neurotoxin of Clostridium botulinum, useful
PT for developing vaccines against neurotoxins of Clostridium botulinum
PT and for identifying protective antibodies
XX
PS Claim 1; Column 25; 14pp; English.

applicant

XX
CC The present sequence represents a fragment of the Clostridium botulinum
CC neurotoxin (BoNT). It was produced by amplifying overlapping fragments
CC of the BoNT gene. The amplified fragments were cloned expressed to
CC identify immunogenic polypeptides which are capable of giving rise to
CC protective antibodies. The BoNT polypeptide fragment are useful as
CC vaccines, for immunizing against botulism, and as diagnostic agents
CC to identify protective antibodies.
XX
SQ Sequence 207 AA;

Query Match 100.0%; Score 1071; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.2e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          |||||||||||||||||||||||||||
Db     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
```

RESULT 2
AAY77141
ID AAY77141 standard; Protein; 407 AA.
XX
AC AAY77141;
XX
DT 08-MAY-2000 (first entry)
XX
DE Native botulinum neurotoxin serotype A (BoNTA) N-terminal fragment (Hn).
XX
KW Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW N-terminal fragment; Hn; Venezuelan equine encephalitis virus replicon;
KW VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS Clostridium botulinum.
XX
PN WO200002524-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US15570.
XX
PR 10-JUL-1998; 98US-0092416.
PR 12-MAY-1999; 99US-0133870.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX
DR WPI; 2000-160827/14.
DR N-PSDB; AAZ87219.
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
PS Example 3; Page 51; 54pp; English.
XX
CC The invention relates to novel vaccines that induce a protective immune
CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC serotypes A-G. In preferred embodiments of the invention, the vector is a
CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC vector results in the production of large amounts of a protein encoded by
CC a sequence cloned into the replicon. The constructs are used to produce
CC vaccines against botulism. The proteins can also be used as diagnostic
CC tools for the diagnosis of botulism. The transformed host cells can be
CC used to analyse the effectiveness of drugs and agents which inhibit toxin
CC effects. The vaccine currently used against botulism is dangerous
CC and expensive to produce, and contains formalin, which is very painful
CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC the 7 serotypes are represented in the formulation. The novel vaccine
CC of overcomes these problems, as it is easily purified, and available in
CC large quantities. It is also expressed in the lymph nodes for a better
CC immune response. The present sequence represents the native BoNTA heavy
CC chain N-terminal fragment (Hn) used in an exemplification of the present
CC invention.
XX
SQ Sequence 407 AA;

Query Match 100.0%; Score 1071; DB 21; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.2e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 61

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 121

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 181

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          |||||||||||||||||||||||||||
Db     182 PYIGPALNIGNMLYKDDFVGALIFSGA 208
```

RESULT 3
AAB04168
ID AAB04168 standard; Protein; 413 AA.
XX
AC AAB04168;
XX
DT 11-APR-2001 (first entry)
XX
DE Botulism toxin heavy chain N-terminal sequence (serotype A).
XX
KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; immune response; vaccine; bacterium;
KW infection.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN WO200067700-A2.
XX
PD 16-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
XX

PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX
DR WPI; 2001-016048/02.
DR N-PSDB; AAA54590.
XX
PT New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulism
XX
PS Disclosure; Fig 11b; 73pp; English.
XX
CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.
XX
SQ Sequence 413 AA;

Query Match 100.0%; Score 1071; DB 22; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.3e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
Db 8 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 67

Qy 61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
Db 68 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 127

Qy 121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
Db 128 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 187

Qy 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
Db 188 PYIGPALNIGNMLYKDDFVGALIFSGA 214

RESULT 4
ABG69074
ID ABG69074 standard; Protein; 413 AA.
XX
AC ABG69074;
XX
DT 07-OCT-2002 (first entry)
XX
DE Botulinum neurotoxin light chain polypeptide #8.
XX
KW Botulinum neurotoxin light chain; BoNT LC; botulism; dystonia; pain;
KW spasticity; ocular motility; facial dyskinesia; stiff-person syndrome;
KW bladder dysfunction; segmental myoclonus; hyperkinetic disorder;
KW cosmetic treatment; facial wrinkle; cerebral palsy; analgesic; relaxant;
KW lower motor neuron hyperactivity; autonomic nerve function; muscular;
KW immunostimulant; antibacterial.
XX
OS Clostridium botulinum.
XX
PN WO200236758-A2.
XX
PD 10-MAY-2002.
XX
PF 06-NOV-2001; 2001WO-US47230.
XX
PR 06-NOV-2000; 2000US-246774P.
PR 20-JUL-2001; 2001US-0910186.
PR 09-AUG-2001; 2001US-311966P.
XX
PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI Smith LA, Jensen M;
XX
DR WPI; 2002-575192/61.
DR N-PSDB; ABK98544.
XX
PT Novel nucleic acid molecule encoding botulinum neurotoxin light chain
PT serotype A, useful for producing the neurotoxin for vaccination against
PT botulism, comprises sequence expressible in host other than Clostridium
PT
XX
PS Disclosure; Page 130-131; 166pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a botulinum
CC neurotoxin light chain (BoNT LC) serotype A, where the DNA has a sequence
CC that is expressible in a host organism other than Clostridium, or has a
CC total A+T content that is less than about 70% The BoNT LC protein is
CC useful in vaccination against botulism, for eliciting protective immunity
CC in a mammal, for treating dystonias, spasticity, pain, ocular motility,
CC facial dyskinesias, stiff-person syndrome, bladder dysfunction, segmental
CC myoclonus, hyperkinetic disorders, cosmetic treatment of facial wrinkles,
CC conditions characterised by hyperactivity of the lower motor neuron, and
CC to control autonomic nerve function or tiptoe-walking due to stiff
CC muscles common in children with cerebral palsy. The sequences are also
CC useful for screening for botulinum neurotoxin inhibitors. This sequence
CC represents a botulinum neurotoxin light chain serotype A protein.
XX
SQ Sequence 413 AA;

Query Match 100.0%; Score 1071; DB 23; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.3e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
Db 8 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 67

Qy 61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
Db 68 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 127

Qy 121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
Db 128 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 187

Qy 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
Db 188 PYIGPALNIGNMLYKDDFVGALIFSGA 214

RESULT 5
AAB36302
ID AAB36302 standard; Protein; 423 AA.
XX
AC AAB36302;
XX
DT 15-FEB-2001 (first entry)
XX
DE C. botulinum BoNT/A neurotoxin heavy chain prototoxin SEQ ID NO:8.
XX
KW Human; procholecystokinin; CCK A receptor; CCK B receptor;
KW pancreatitis; antiinflammatory.
XX

OS Clostridium botulinum.
XX
PN WO200061192-A2.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09142.
XX
PR 08-APR-1999; 99US-0288326.
XX
PA (ALLR) ALLERGAN SALES INC.
XX
PI Steward LE, Sachs G, Aoki KR;
XX
DR WPI; 2000-679416/66.
XX
PT New composition for treating acute pancreatitis, comprises a pancreatic
PT cell surface marker binding element, a translocation element that
PT transfers polypeptide across vesicular membrane, and a therapeutic
PT element -
XX
PS Disclosure; Page 28; 50pp; English.
XX
CC The present invention describes a composition (I) for treating acute
CC pancreatitis. (I) comprises a first element containing a binding element
CC that binds to a pancreatic cell surface marker, a second element
CC containing a translocation element that facilitates polypeptide transfer
CC across the vesicular membrane, and a third element containing a
CC therapeutic element that inhibits enzyme secretion in pancreatic cell
CC cytoplasm. Also described is a method for making a therapeutic
CC polypeptide having a binding element selective for cholecystokinin (CCK)
CC receptor by expressing within a host cell a recombinant chimeric
CC polypeptide comprising an extein containing a therapeutic element and a
CC translocational element, and an intein located to the carboxy terminal
CC of extein having a cysteine, serine or threonine at its amino terminus,
CC and contacting the extein with a synthetic peptide comprising a CCK
CC amino acid sequence containing an amidated phenylalanine at a natural
CC C-terminus, and a cysteine, serine or threonine at its N-terminus, and
CC a nucleophilic reagent able to cause cleavage of the intein to form a
CC peptide bond between the extein C-terminus and synthetic peptide
CC N-terminus through the formation of an activated ester or thio ester
CC intermediate. (I) has antiinflammatory activity and prevents accumulation
CC of pancreatic digestive enzymes, and prevents exocytic fusion of vesicles
CC containing secretory enzymes of pancreas. (I) is useful for treating
CC acute pancreatitis. The present sequence represents the Clostridium
CC botulinum BoNT/A neurotoxin heavy chain prototoxin which is given in
CC the exemplification of the present invention.
XX
SQ Sequence 423 AA;

Query Match 100.0%; Score 1071; DB 21; Length 423;
Best Local Similarity 100.0%; Pred. No. 3.4e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 66

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 126

Qy    121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 186

Qy    181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db    187 PYIGPALNIGNMLYKDDFVGALIFSGA 213
```

RESULT 6
AAY77140

ID AAY77140 standard; Protein; 837 AA.
XX
AC AAY77140;
XX
DT 08-MAY-2000 (first entry)
XX
DE Native botulinum neurotoxin serotype A (BoNTA).
XX
KW Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW Venezuelan equine encephalitis virus replicon;
KW VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS Clostridium botulinum.
XX
FH Key Location/Qualifiers
FT Misc-difference 837
FT /note= "Apparently encoded by GGATGGGGAG AAAGGCCACT G"
XX
PN WO200002524-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US15570.
XX
PR 10-JUL-1998; 98US-0092416.
PR 12-MAY-1999; 99US-0133870.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX
DR WPI; 2000-160827/14.
DR N-PSDB; AAZ87218.
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
PS Example 3; Page 49; 54pp; English.
XX
CC The invention relates to novel vaccines that induce a protective immune
CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC serotypes A-G. In preferred embodiments of the invention, the vector is a
CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC vector results in the production of large amounts of a protein encoded by
CC a sequence cloned into the replicon. The constructs are used to produce
CC vaccines against botulism. The proteins can also be used as diagnostic
CC tools for the diagnosis of botulism. The transformed host cells can be
CC used to analyse the effectiveness of drugs and agents which inhibit toxin
CC effects. The vaccine currently used against botulism is dangerous
CC and expensive to produce, and contains formalin, which is very painful
CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC the 7 serotypes are represented in the formulation. The novel vaccine
CC of overcomes these problems, as it is easily purified, and available in
CC large quantities. It is also expressed in the lymph nodes for a better
CC immune response. The present sequence represents the native BoNTA heavy
CC chain used in an exemplification of the present invention.
XX
SQ Sequence 837 AA;

Query Match 100.0%; Score 1071; DB 21; Length 837;
Best Local Similarity 100.0%; Pred. No. 8.9e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 61

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      62 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 121

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 181

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
           ||||||||||||||||||||||||||||
Db     182 PYIGPALNIGNMLYKDDFVGALIFSGA 208


RESULT 7
ABG69075
ID   ABG69075 standard; Protein; 861 AA.
XX
AC   ABG69075;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Botulinum neurotoxin light chain polypeptide #9a.
XX
KW   Botulinum neurotoxin light chain; BoNT LC; botulism; dystonia; pain;
KW   spasticity; ocular motility; facial dyskinesia; stiff-person syndrome;
KW   bladder dysfunction; segmental myoclonus; hyperkinetic disorder;
KW   cosmetic treatment; facial wrinkle; cerebral palsy; analgesic; relaxant;
KW   lower motor neuron hyperactivity; autonomic nerve function; muscular;
KW   immunostimulant; antibacterial.
XX
OS   Clostridium botulinum.
XX
PN   WO200236758-A2.
XX
PD   10-MAY-2002.
XX
PF   06-NOV-2001; 2001WO-US47230.
XX
PR   06-NOV-2000; 2000US-246774P.
PR   20-JUL-2001; 2001US-0910186.
PR   09-AUG-2001; 2001US-311966P.
XX
PA   (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI   Smith LA,  Jensen M;
XX
DR   WPI; 2002-575192/61.
DR   N-PSDB; ABK98545.
XX
PT   Novel nucleic acid molecule encoding botulinum neurotoxin light chain
PT   serotype A, useful for producing the neurotoxin for vaccination against
PT   botulism, comprises sequence expressible in host other than Clostridium
PT    -
XX
PS   Claim 13; Page 133-134; 166pp; English.
XX
CC   The invention relates to a nucleic acid molecule encoding a botulinum
CC   neurotoxin light chain (BoNT LC) serotype A, where the DNA has a sequence
CC   that is expressible in a host organism other than Clostridium, or has a
CC   total A+T content that is less than about 70% The BoNT LC protein is
CC   useful in vaccination against botulism, for eliciting protective immunity
CC   in a mammal, for treating dystonias, spasticity, pain, ocular motility,
CC   facial dyskinesias, stiff-person syndrome, bladder dysfunction, segmental
CC   myoclonus, hyperkinetic disorders, cosmetic treatment of facial wrinkles,
CC   conditions characterised by hyperactivity of the lower motor neuron, and
CC   to control autonomic nerve function or tiptoe-walking due to stiff
CC   muscles common in children with cerebral palsy. The sequences are also
CC   useful for screening for botulinum neurotoxin inhibitors. This sequence
CC   represents a botulinum neurotoxin light chain serotype A protein.
XX
SQ   Sequence   861 AA;

  Query Match           100.0%;  Score 1071;  DB 23;  Length 861;
  Best Local Similarity  100.0%;  Pred. No. 9.3e-99;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     456 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 515

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     516 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 575

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     576 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 635

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
           ||||||||||||||||||||||||||||
Db     636 PYIGPALNIGNMLYKDDFVGALIFSGA 662


RESULT 8
AAW56019
ID   AAW56019 standard; Protein; 871 AA.
XX
AC   AAW56019;
XX
DT   27-JUL-1998  (first entry)
XX
DE   Recombinant botulinum neurotoxin type A LH423/A (Q2E,N26K,A27Y).
XX
KW   Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW   immunogen; detection; tetanus; non-toxic; toxin.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO9807864-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-GB02273.
XX
PR   13-DEC-1996;   96GB-0025996.
PR   23-AUG-1996;   96GB-0017671.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA   (SPEY-) SPEYWOOD LAB LTD.
XX
PI   Foster KA,  Quinn CP,  Shone CC;
XX
DR   WPI; 1998-169168/15.
DR   N-PSDB; AAV26291.
XX
PT   Recombinant neurotoxin polypeptides - used to develop therapeutic
PT   agents, immunogens or as non-toxic standards for the detection of
PT   neurotoxins
XX
PS   Example 1; Page 108-111; 137pp; English.
XX
CC   The present sequence represents a recombinant neurotoxin protein from
CC   the present invention. The present invention describes recombinant
CC   neurotoxin proteins which comprise a first and second domain, where
CC   the first domain is adapted to cleave one or more vesicle or
CC   plasma-membrane associated proteins essential to exocytosis, and where
CC   the second domain is adapted: (a) to translocate the protein into a
CC   cell; (b) to increase the solubility of the protein compared to the
CC   solubility of the first domain on its own, or (c) both to translocate
CC   the protein into a cell and to increase the solubility of the protein
CC   compared to the solubility of the first domain on its own, the protein
CC   being free of clostridial neurotoxin (CN) and free of CN precursor that
CC   can be converted into toxin by proteolytic action. The recombinant
CC   proteins can be used as therapeutic agents for targeting cells
CC   expressing a relevant substrate. The products can also be used as
CC   immunogens and as non-toxic standards for the assessment and development
CC   of in vitro assays for the detection of functional botulinum or tetanus
```

```
CC   neurotoxins either in foodstuffs or in environmental samples.
XX
SQ   Sequence   871 AA;

  Query Match           100.0%;  Score 1071;  DB 19;  Length 871;
  Best Local Similarity  100.0%;  Pred. No. 9.4e-99;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 514

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 574

Qy    121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 634

Qy    181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db    635 PYIGPALNIGNMLYKDDFVGALIFSGA 661


RESULT 9
AAW56007
ID   AAW56007 standard; Protein; 871 AA.
XX
AC   AAW56007;
XX
DT   27-JUL-1998  (first entry)
XX
DE   Recombinant botulinum neurotoxin type A LH423/A.
XX
KW   Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW   immunogen; detection; tetanus; non-toxic; toxin.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO9807864-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-GB02273.
XX
PR   13-DEC-1996;   96GB-0025996.
PR   23-AUG-1996;   96GB-0017671.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA   (SPEY-) SPEYWOOD LAB LTD.
XX
PI   Foster KA,  Quinn CP,  Shone CC;
XX
DR   WPI; 1998-169168/15.
DR   N-PSDB; AAV26279.
XX
PT   Recombinant neurotoxin polypeptides - used to develop therapeutic
PT   agents, immunogens or as non-toxic standards for the detection of
PT   neurotoxins
XX
PS   Example 1; Page 33-35; 137pp; English.
XX
CC   The present sequence represents a recombinant neurotoxin protein from
CC   the present invention. The present invention describes recombinant
CC   neurotoxin proteins which comprise a first and second domain, where
CC   the first domain is adapted to cleave one or more vesicle or
CC   plasma-membrane associated proteins essential to exocytosis, and where
CC   the second domain is adapted: (a) to translocate the protein into a
CC   cell; (b) to increase the solubility of the protein compared to the
CC   solubility of the first domain on its own, or (c) both to translocate
CC   the protein into a cell and to increase the solubility of the protein
CC   compared to the solubility of the first domain on its own, the protein
CC   being free of clostridial neurotoxin (CN) and free of CN precursor that
CC   can be converted into toxin by proteolytic action. The recombinant
CC   proteins can be used as therapeutic agents for targeting cells
CC   expressing a relevant substrate. The products can also be used as
CC   immunogens and as non-toxic standards for the assessment and development
CC   of in vitro assays for the detection of functional botulinum or tetanus
CC   neurotoxins either in foodstuffs or in environmental samples.
XX
SQ   Sequence   871 AA;

  Query Match           100.0%;  Score 1071;  DB 19;  Length 871;
  Best Local Similarity  100.0%;  Pred. No. 9.4e-99;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 514

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 574

Qy    121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 634

Qy    181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db    635 PYIGPALNIGNMLYKDDFVGALIFSGA 661


RESULT 10
AAW56008
ID   AAW56008 standard; Protein; 871 AA.
XX
AC   AAW56008;
XX
DT   27-JUL-1998  (first entry)
XX
DE   Botulinum neurotoxin type A BoNT/A.
XX
KW   Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW   immunogen; detection; tetanus; non-toxic; toxin.
XX
OS   Clostridium botulinum.
XX
PN   WO9807864-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-GB02273.
XX
PR   13-DEC-1996;   96GB-0025996.
PR   23-AUG-1996;   96GB-0017671.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA   (SPEY-) SPEYWOOD LAB LTD.
XX
PI   Foster KA,  Quinn CP,  Shone CC;
XX
DR   WPI; 1998-169168/15.
DR   N-PSDB; AAV26280.
XX
PT   Recombinant neurotoxin polypeptides - used to develop therapeutic
PT   agents, immunogens or as non-toxic standards for the detection of
PT   neurotoxins
XX
PS   Disclosure; Page 52-54; 137pp; English.
XX
CC   The present sequence represents botulinum neurotoxin type A from
CC   the present invention. The present invention describes recombinant
CC   neurotoxin proteins which comprise a first and second domain, where
```

```
CC   the first domain is adapted to cleave one or more vesicle or
CC   plasma-membrane associated proteins essential to exocytosis, and where
CC   the second domain is adapted: (a) to translocate the protein into a
CC   cell; (b) to increase the solubility of the protein compared to the
CC   solubility of the first domain on its own, or (c) both to translocate
CC   the protein into a cell and to increase the solubility of the protein
CC   compared to the solubility of the first domain on its own, the protein
CC   being free of clostridial neurotoxin (CN) and free of CN precursor that
CC   can be converted into toxin by proteolytic action. The recombinant
CC   proteins can be used as therapeutic agents for targeting cells
CC   expressing a relevant substrate. The products can also be used as
CC   immunogens and as non-toxic standards for the assessment and development
CC   of in vitro assays for the detection of functional botulinum or tetanus
CC   neurotoxins either in foodstuffs or in environmental samples.
XX
SQ   Sequence   871 AA;

  Query Match            100.0%;  Score 1071;  DB 19;  Length 871;
  Best Local Similarity  100.0%;  Pred. No. 9.4e-99;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 514

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 574

Qy    121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 634

Qy    181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          |||||||||||||||||||||||||||
Db    635 PYIGPALNIGNMLYKDDFVGALIFSGA 661


RESULT 11
AAW56016
ID   AAW56016 standard; Protein; 873 AA.
XX
AC   AAW56016;
XX
DT   27-JUL-1998  (first entry)
XX
DE   Recombinant botulinum neurotoxin type A 2LH423/A (Q2E,N26K,A27Y).
XX
KW   Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW   immunogen; detection; tetanus; non-toxic; toxin.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO9807864-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-GB02273.
XX
PR   13-DEC-1996;   96GB-0025996.
PR   23-AUG-1996;   96GB-0017671.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA   (SPEY-) SPEYWOOD LAB LTD.
XX
PI   Foster KA,  Quinn CP,  Shone CC;
XX
DR   WPI; 1998-169168/15.
DR   N-PSDB; AAV26288.
XX
PT   Recombinant neurotoxin polypeptides - used to develop therapeutic
PT   agents, immunogens or as non-toxic standards for the detection of
PT   neurotoxins
XX
PS   Example 1; Page 45-48; 137pp; English.
XX
CC   The present sequence represents a recombinant neurotoxin protein from
CC   the present invention. The present invention describes recombinant
CC   neurotoxin proteins which comprise a first and second domain, where
CC   the first domain is adapted to cleave one or more vesicle or
CC   plasma-membrane associated proteins essential to exocytosis, and where
CC   the second domain is adapted: (a) to translocate the protein into a
CC   cell; (b) to increase the solubility of the protein compared to the
CC   solubility of the first domain on its own, or (c) both to translocate
CC   the protein into a cell and to increase the solubility of the protein
CC   compared to the solubility of the first domain on its own, the protein
CC   being free of clostridial neurotoxin (CN) and free of CN precursor that
CC   can be converted into toxin by proteolytic action. The recombinant
CC   proteins can be used as therapeutic agents for targeting cells
CC   expressing a relevant substrate. The products can also be used as
CC   immunogens and as non-toxic standards for the assessment and development
CC   of in vitro assays for the detection of functional botulinum or tetanus
CC   neurotoxins either in foodstuffs or in environmental samples.
XX
SQ   Sequence   873 AA;

  Query Match            100.0%;  Score 1071;  DB 19;  Length 873;
  Best Local Similarity  100.0%;  Pred. No. 9.5e-99;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 516

Qy     61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    517 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 576

Qy    121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    577 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 636

Qy    181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          |||||||||||||||||||||||||||
Db    637 PYIGPALNIGNMLYKDDFVGALIFSGA 663


RESULT 12
AAW56009
ID   AAW56009 standard; Protein; 875 AA.
XX
AC   AAW56009;
XX
DT   27-JUL-1998  (first entry)
XX
DE   Recombinant botulinum neurotoxin type A L/4H423/A.
XX
KW   Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW   immunogen; detection; tetanus; non-toxic; toxin.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO9807864-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-GB02273.
XX
PR   13-DEC-1996;   96GB-0025996.
PR   23-AUG-1996;   96GB-0017671.
XX
PA   (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA   (SPEY-) SPEYWOOD LAB LTD.
XX
```

PI Foster KA, Quinn CP, Shone CC;
XX
DR WPI; 1998-169168/15.
DR N-PSDB; AAV26281.
XX
PT Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
XX
PS Example 1; Page 58-60; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 875 AA;

Query Match 100.0%; Score 1071; DB 19; Length 875;
Best Local Similarity 100.0%; Pred. No. 9.5e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     459 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 518

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     519 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 578

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     579 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 638

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db     639 PYIGPALNIGNMLYKDDFVGALIFSGA 665
```

RESULT 13
AAW56010
ID AAW56010 standard; Protein; 878 AA.
XX
AC AAW56010;
XX
DT 27-JUL-1998 (first entry)
XX
DE Recombinant botulinum neurotoxin type A LFXa/3H423/A.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN WO9807864-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-GB02273.
XX
PR 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
XX
PI Foster KA, Quinn CP, Shone CC;
XX
DR WPI; 1998-169168/15.
DR N-PSDB; AAV26282.
XX
PT Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
XX
PS Example 1; Page 64-66; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 878 AA;

Query Match 100.0%; Score 1071; DB 19; Length 878;
Best Local Similarity 100.0%; Pred. No. 9.5e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 521

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 581

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 641

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db     642 PYIGPALNIGNMLYKDDFVGALIFSGA 668
```

RESULT 14
AAW56015
ID AAW56015 standard; Protein; 894 AA.
XX
AC AAW56015;
XX
DT 27-JUL-1998 (first entry)
XX
DE Recombinant botulinum neurotoxin type A 23LH423/A (Q2E,N26K,A27Y).
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.

OS Clostridium botulinum.
XX
PN WO9807864-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-GB02273.
XX
PR 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
XX
PI Foster KA, Quinn CP, Shone CC;
XX
DR WPI; 1998-169168/15.
DR N-PSDB; AAV26287.
XX
PT Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
XX
PS Example 1; Page 39-42; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 894 AA;

Query Match 100.0%; Score 1071; DB 19; Length 894;
Best Local Similarity 100.0%; Pred. No. 9.8e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 537

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 597

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 657

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db     658 PYIGPALNIGNMLYKDDFVGALIFSGA 684
```

RESULT 15
AAW56012
ID AAW56012 standard; Protein; 907 AA.
XX
AC AAW56012;
XX
DT 27-JUL-1998 (first entry)
XX
DE Recombinant botulinum neurotoxin type A LFXa/3H423/A-CtxA 14.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN WO9807864-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-GB02273.
XX
PR 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
XX
PI Foster KA, Quinn CP, Shone CC;
XX
DR WPI; 1998-169168/15.
DR N-PSDB; AAV26284.
XX
PT Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
XX
PS Example 1; Page 77-79; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 1071; DB 19; Length 907;
Best Local Similarity 100.0%; Pred. No. 1e-98;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 521

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 581

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 641

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db     642 PYIGPALNIGNMLYKDDFVGALIFSGA 668
```

Search completed: March 13, 2003, 11:39:09
Job time : 41.7752 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:38:07 ; Search time 16.1066 Seconds
(without alignments)
1235.505 Million cell updates/sec

Title: US-09-917-791-21
Perfect score: 1071
Sequence: 1 IKVNNWDLFFSPSEDNFTND.....NIGNMLYKDDFVGALIFSGA 207

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1071	100.0	1296	1	BTCLAB	bontoxilysin (EC 3
2	918	85.7	1296	2	I40645	botulinum neurotox
3	407.5	38.0	1268	2	S33411	botulinum neurotox
4	377	35.2	1297	2	S39791	neurotoxin - Clost
5	374.5	35.0	1251	2	JH0256	botulinum neurotox
6	368.5	34.4	1252	2	S21178	botulinum neurotox
7	366.5	34.2	1291	1	A48940	bontoxilysin (EC 3
8	357.5	33.4	1291	2	I40631	non-proteolytic bo
9	316	29.5	1274	2	I40813	neurotoxin type F
10	312	29.1	1315	1	BTCLTN	tentoxilysin (EC 3
11	303.5	28.3	1285	2	S70582	botulinum neurotox
12	290.5	27.1	1276	2	S11455	botulinum neurotox
13	263	24.6	1291	2	S46431	botulinum neurotox
14	263	24.6	1291	2	A49777	botulinum neurotox
15	107	10.0	246	2	A84977	hypothetical prote
16	106	9.9	1196	2	JQ1467	toxin, nontoxic co
17	106	9.9	1196	2	S46430	botulinum neurotox
18	105	9.8	2401	2	T28676	rhoptry protein -
19	104	9.7	1193	2	JC4901	nontoxic-nonhemagg
20	99	9.2	1193	2	S68218	botulinum neurotox
21	98.5	9.2	4550	2	T18440	hypothetical prote
22	98	9.2	345	2	S55969	H+-exporting ATPas
23	98	9.2	2269	2	T28677	rhoptry protein -
24	97.5	9.1	356	2	D97732	hypothetical prote
25	97	9.1	1008	2	T18508	hypothetical prote
26	95.5	8.9	1032	2	S53571	hypothetical prote
27	93.5	8.7	355	2	H71730	hypothetical prote
28	92	8.6	1162	2	A47708	progenitor toxin n
29	91	8.5	635	2	B82255	heat shock protein
30	91	8.5	1162	2	I40817	botulinum toxin no
31	90.5	8.5	862	2	C97343	hypothetical prote
32	90.5	8.5	1144	2	A36968	P1-like adhesin pr
33	90.5	8.5	1186	2	C64588	cag pathogenicity
34	89.5	8.4	1203	2	D86625	subunit A of ATP-d
35	89	8.3	178	2	S75511	hypothetical prote
36	89	8.3	1146	2	E70204	hypothetical prote
37	89	8.3	1215	2	S60904	hypothetical prote
38	88.5	8.3	556	2	AD1770	arginyl tRNA synth
39	87.5	8.2	419	2	G90213	hypothetical prote
40	87.5	8.2	520	2	C81273	hypothetical prote
41	87.5	8.2	1254	2	T18425	hypothetical prote
42	87	8.1	353	2	T15724	hypothetical prote
43	87	8.1	1516	2	E71619	RAD2 endonuclease
44	86.5	8.1	556	2	AI1394	arginyl tRNA synth
45	86.5	8.1	2510	2	T28160	hypothetical prote

ALIGNMENTS

RESULT 1
BTCLAB
bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type A
C;Species: Clostridium botulinum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: A35294
A;Molecule type: DNA
A;Residues: 1-1296 <BIN>
A;Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
A;Experimental source: strain 62A, subtype A
R;Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990
A;Title: The complete amino acid sequence of the Clostridium botulinum type A neuroto
A;Reference number: S09492; MUID:90235864; PMID:2185020
A;Accession: S09492
A;Molecule type: DNA
A;Residues: 1,'Q',3-26,'V',28-1296 <THO>
A;Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382
A;Experimental source: NCTC 2916
R;Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone
A;Reference number: S67988; MUID:96096783; PMID:8521962
A;Accession: S68220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <FUJ>
A;Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
R;Betley, M.J.; Somers, E.; DasGupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A;Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t
A;Reference number: A33401; MUID:89350959; PMID:2669749
A;Accession: A33401
A;Molecule type: DNA
A;Residues: 1-35 <BET>
A;Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
R;Gimenez, J.A.; DasGupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A;Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
A;Reference number: A53884; MUID:94000342; PMID:8397793
A;Accession: A53884
A;Status: preliminary
A;Molecule type: protein
A;Residues: 867-880;1148-1217,'Y',1219 <GIM>
A;Experimental source: strain Hall
A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after

```
R;DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and arou
A;Reference number: A60025; MUID:91120847; PMID:2126206
A;Accession: A60025
A;Molecule type: protein
A;Residues: 2-6;445-453,'X',455-457 <DAS1>
R;DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A;Reference number: A27000
A;Accession: A27000
A;Molecule type: protein
A;Residues: 2-47 <DAS2>
R;Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A;Title: Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.
A;Reference number: A49708; MUID:94124495; PMID:8294407
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
C;Genetics:
A;Gene: atx; botA
C;Function:
A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
F;2-444/Product: bontoxilysin A light chain #status experimental <LGHT>
F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
F;223,227/Binding site: zinc (His) #status predicted
F;224/Active site: Glu #status predicted

  Query Match             100.0%;  Score 1071;  DB 1;  Length 1296;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-77;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     455 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 514

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     515 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 574

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     575 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 634

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||||||||||||||||||||||||||||
Db     635 PYIGPALNIGNMLYKDDFVGALIFSGA 661


RESULT 2
I40645
botulinum neurotoxin type A - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40645
R;Willems, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A;Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A;Reference number: I40645; MUID:94143603; PMID:8310180
A;Accession: I40645
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1296 <RES>
A;Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             85.7%;  Score 918;  DB 2;  Length 1296;
  Best Local Similarity   85.4%;  Pred. No. 2.4e-65;
  Matches  176;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
          |||||||||||||||||||:| ||||:|||||||||||||||||||||:|||||||
Db     455 IKVNNWDLFFSPSEDNFTNDLDKVEEITADTNIEAAEENISLDLIQQYYLTFDFDNEPEN 514

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
          ||||||||||||||| ||||||||||||||||||||||||||||||| ||| ||||  |
Db     515 ISIENLSSDIIGQLEPMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGDSRIILTNSAEE 574

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
          ||| |:  |||||| ||||:||| || ||| | |:||||||||:||:| |||||||||:
Db     575 ALLKPNVAYTFFSSKYVKKINKAVEAFMFLNWAEELVYDFTDETNEVTTMDKIADITIIV 634

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSG 206
          |||||||||||||| | :|| |:||:|
Db     635 PYIGPALNIGNMLSKGEFVEAIIFTG 660


RESULT 3
S33411
botulinum neurotoxin type F - Clostridium barati
C;Species: Clostridium barati
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S33411; S31860
R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P
FEMS Microbiol. Lett. 108, 175-182, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neuroto
A;Reference number: S33411; MUID:93252228; PMID:8486245
A;Accession: S33411
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1268 <THO>
A;Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             38.0%;  Score 407.5;  DB 2;  Length 1268;
  Best Local Similarity   44.0%;  Pred. No. 1.8e-24;
  Matches   95;  Conservative   39;  Mismatches   61;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
          ||||| ||||  || ::  | :|  :||  || |       :||   :  | :| |
Db     436 IKVNNRDLFFVASESSYNENGINSPKEI-DDTTITNNNYKKNLD---EVILDYNSD---- 487

Qy      60 NISIENLSSDIIGQL----ELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
            :| |||| ::          :|  :   ||    |:| :||  :| || ||:   |:|
Db     488 --AIPNLSSRLLNTTAQNDSYVPKYD--SNGTSEIKEYTVDKLNVFFYLYAQKAPEGESA 543

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
          |:||:||| |||: |:|||||||::  |||  :||:|: |::|:: ||| | :: || |
Db     544 ISLTSSVNTALLDASKVYTFFSSDFINTVNKPVQAALFISWIQQVINDFTTEATQKSTID 603

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
          |||||::|:||:| |||||| : | :|  |:   ||
Db     604 KIADISLIVPYVGLALNIGNEVQKGNFKEAIELLGA 639


RESULT 4
S39791
neurotoxin - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C;Accession: S39791
R;Campbell, K.; Collins, M.D.; East, A.K.
Biochim. Biophys. Acta 1216, 487-491, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridiu
A;Reference number: S39791; MUID:94092745; PMID:8268233
A;Accession: S39791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1297 <CAM>
A;Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276
C;Superfamily: tetanus toxin
```

```
C;Keywords: neurotoxin

  Query Match             35.2%;  Score 377;  DB 2;  Length 1297;
  Best Local Similarity   41.3%;  Pred. No. 5.2e-22;
  Matches   88;  Conservative   34;  Mismatches   75;  Indels   16;  Gaps    3;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLD-LIQQYYLTFNFDNEPE 59
Db     451 IIVNNEDLFFIANKDSFSKDLAKAETIAYNTQNNTIENNFSIDQLILDNDLSSGIDLPNE 510

Qy      60 NISIENLSSDI-----IGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIAL 114
Db     511 NTEPFTNFDDIDIPVYIKQSAL----------KKIFVDGDSLFEYLHAQTFPSNIENLQL 560

Qy     115 TNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA 174
Db     561 TNSLNDALRNNNKVYTFFSTNLVEKANTVVGASLFVNWVKGVIDDFTSESTQKSTIDKVS 620

Qy     175 DITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
Db     621 DVSIIIPYIGPALNVGNETAKENFKNAFEIGGA 653


RESULT 5
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C;Accession: JH0256; S16145
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum type
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0256
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-27,'E',29-1251 <POU>
A;Cross-references: EMBL:X62088; NID:g40379
A;Experimental source: strains ATCC 43181 and ATCC 43755
R;Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.
J. Gen. Microbiol. 137, 519-525, 1991
A;Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxi
A;Reference number: S16145; MUID:91237316; PMID:2033376
A;Accession: S16145
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229,'M',231-252 <FUJ>
A;Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408
A;Experimental source: strain BL6340
C;Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F;423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F;412-426/Disulfide bonds: #status predicted

  Query Match             35.0%;  Score 374.5;  DB 2;  Length 1251;
  Best Local Similarity   38.9%;  Pred. No. 7.9e-22;
  Matches   84;  Conservative   51;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
Db     427 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 478

Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
Db     479 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 534

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
Db     535 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQKSTVD 594

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
Db     595 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 630


RESULT 6
S21178
botulinum neurotoxin type E precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
C;Accession: S21178; S48107; JH0257; B35294; A60027; S18111
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992
A;Title: The complete amino acid sequence of the Clostridium botulinum type-E neuroto
A;Reference number: S21178; MUID:92174922; PMID:1541280
A;Accession: S21178
A;Molecule type: DNA
A;Residues: 1-1252 <WHE>
A;Cross-references: EMBL:X62683; NID:g40397; PIDN:CAA44558.1; PID:g40398
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and sp ific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48107
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 616-982 <CAM>
A;Cross-references: EMBL:X70815; NID:g407786; PIDN:CAA50146.1; PID:g407787
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum t
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0257
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-176,'R',178-197,'C',199-339,'R',341-772,'I',774-962,'FE',965-966,'R',96
A;Cross-references: EMBL:X62089; NID:g40393; PIDN:CAA43999.1; PID:g40394
A;Experimental source: strain Beluga
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: B35294
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-176,'R',178-252 <BIN>
A;Experimental source: strain Beluga
R;Gimenez, J.A.; DasGupta, B.R.
Biochimie 72, 213-217, 1990
A;Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reve he
A;Reference number: A60027; MUID:90344918; PMID:2116911
A;Accession: A60027
A;Molecule type: protein
A;Residues: 420-427 <GIM>
A;Experimental source: strain Beluga
A;Note: this fragment was generated by proteolysis with Lys-C rather than with trypsi
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F;423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
F;412-426/Disulfide bonds: #status predicted

  Query Match             34.4%;  Score 368.5;  DB 2;  Length 1252;
  Best Local Similarity   38.4%;  Pred. No. 2.4e-21;
  Matches   83;  Conservative   51;  Mismatches   61;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
Db     427 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 478
```

```
Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           : |  :| | |:  |      :|  :   ||    :::::::  :| || ||:   |::
Db     479 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 534

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           : ||:|:: |||   ::||||||:::  |||  :||:|: |::|:: ||| | :: || |
Db     535 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVD 594

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||:|::|||| ||||||   | :|  ||   ||
Db     595 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 630
```

RESULT 7
A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type B (BoNT/B)
C;Species: Clostridium botulinum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the typ
A;Reference number: A48940; MUID:92384550; PMID:1514783
A;Accession: A48940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <WHE>
A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A;Experimental source: type B, Danish
A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publica
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-994 <CAM>
A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A;Experimental source: proteolytic type B, strain NCTC 7273
R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A;Description: Partial amino acid sequence of botulinum neurotoxin type B and comparisic
A;Reference number: S21575
A;Accession: S21575
A;Molecule type: DNA
A;Residues: 36-217,'G',219-224,'S',226-246 <SZA>
A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K
J. Biol. Chem. 267, 14721-14729, 1992
A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus to
A;Reference number: A42871; MUID:92340509; PMID:1634516
A;Accession: A42871
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-313,'S',315-451 <KUR>
A;Experimental source: strain Okra
A;Note: sequence extracted from NCBI backbone (NCBIP:109365)
R;DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
A;Reference number: S07155; MUID:89000987; PMID:3139097
A;Accession: S07155
A;Molecule type: protein
A;Residues: 2-29,'M',31-45 <DAS>
A;Accession: S08562
A;Molecule type: protein
A;Residues: 442-463,'R',465-467 <DA2>
R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A;Reference number: S07128; MUID:85197963; PMID:3888113
A;Accession: S07128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-16 <SCH1>
A;Accession: S08573
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-17 <SCH2>
A;Accession: S08574
A;Status: preliminary
A;Molecule type: protein
A;Residues: 442-459 <SCH3>
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinerg[redacted]gy
C;Genetics:
A;Gene: bont/b
C;Function:
A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

  Query Match   34.2%; Score 366.5; DB 1; Length 1291;
  Best Local Similarity 39.4%; Pred. No. 3.6e-21;
  Matches   87; Conservative   39; Mismatches   62; Indels   33; Gaps   4;

```
Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           | |:| ||||   :::|::||:| | |  :|             |  |:  :|   | |
Db     447 IDVDNEDLFFIADKNSFSDDLSKNERIEYNT-------------QSNYIENDF---PIN 489

Qy      61 ISIENLSSDIIGQLEL-------------MPNIERFPNGKKYELDKYTMFHYLRAQEFE 106
             |  | :|:| ::||             :|  |: |  ||   |: |:| || :| |
Db     490 ELI--LDTDLISKIELPSENTESLTDFNVDVPVYEKQPAIKKIFTDENTIFQYLYSQTFP 547

Qy     107 HGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSE 166
                 |:||:| ::|||  ::||:||| ||:|  ||  || :| |||:|:| ||  | ::
Db     548 LDIRDISLTSSFDDALLFSNKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVNDFVIEANK 607

Qy     167 VSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
             :| ||||||::|:|||| |||:||   | :|  |   :||
Db     608 SNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFEIAGA 648
```

RESULT 8
I40631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40631; S48103; S48104; S36015
R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botul
A;Reference number: I40631; MUID:94122659; PMID:7764370
A;Accession: I40631
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1291 <RES>
A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761,'E',763-841,'M',843,'T',845,'N',847-994 <CAM1>

```
A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A;Experimental source: non-proteolytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843,'T',845,'N',847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C;Comment: Botulinum neurotoxin type B in these strains may posses a capable catalytic s
C;Genetics:
A;Gene: bont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

  Query Match             33.4%;  Score 357.5;  DB 2;  Length 1291;
  Best Local Similarity   38.5%;  Pred. No. 1.9e-20;
  Matches   85;  Conservative   41;  Mismatches   62;  Indels   33;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
Db     447 IDVDNENLFFIADKNSFSDDLSKNERVEYNT--------------QNNYIGNDF---PIN 489

Qy      61 ISIENLSSDIIGQLEL--------------MPNIERFPNGKKYELDKYTMFHYLRAQEFE 106
Db     490 ELI--LDTDLISKIELPSENTESLTDFNVDVPVYEKQPAIKKVFTDENTIFQYLYSQTFP 547

Qy     107 HGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSE 166
Db     548 LNIRDISLTSSFDDALLVSSKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVDDFVIEANK 607

Qy     167 VSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
Db     608 SSTMDKIADISLIVPYIGLALNVGDETAKGNFESAFEIAGS 648


RESULT 9
I40813
neurotoxin type F - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40813; S48108
R;East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E
FEMS Microbiol. Lett. 96, 225-230, 1992
A;Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A;Reference number: I40644
A;Accession: I40813
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1274 <RES>
A;Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48108
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 634-1002 <CAM>
A;Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             29.5%;  Score 316;  DB 2;  Length 1274;
  Best Local Similarity   36.2%;  Pred. No. 3.9e-17;
  Matches   79;  Conservative   51;  Mismatches   62;  Indels   26;  Gaps   11;

Qy       1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
Db     446 IRVNNSELFFVASESSYNENDINTPKEIDDTTNLNNNYRN-NLD---EVILDYNSQTIPQ 501

Qy      60 NISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSRIALT 115
Db     502 -ISNRTLNT-LVQDNSYVPRYD--SNGTSEIEEYDVVDFNVFFYLHAQKVPEGETNISLT 557

Qy     116 NSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIAD 175
Db     558 SSIDTALLEESK-DIFFSSEFIDTINKPVNAALFIDWISKVIRDFTTEATQKSTVDKIAD 616

Qy     176 ITIIIPYIGPALNI------GNMLYKDDF----VGALI 203
Db     617 ISLIVPYVGLALNIIIEAEKGN--FEEAFELLGVGILL 652


RESULT 10
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; [redacted]el,
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A;Experimental source: strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-termin
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation; epitope region
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotox
A;Reference number: S69348; MUID:95262688; PMID:7744050
A;Accession: S69348
A;Molecule type: protein
```

```
A;Residues: 2-31 <DEF>
C;Comment: The source of this protein was an extrachromosomal plasmid.
C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglic
C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C;Function:
A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-457/Product: tentoxylysin light chain (fragment A) #status predicted <TTL>
F;461-1315/Product: tentoxylysin heavy chain (fragment B.C) #status experimental <TTH>
F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F;233,237/Binding site: zinc (His) #status predicted
F;234/Active site: Glu #status predicted

  Query Match             29.1%;  Score 312;  DB 1;  Length 1315;
  Best Local Similarity   33.2%;  Pred. No. 8.6e-17;
  Matches   74;  Conservative   44;  Mismatches   67;  Indels   38;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||: | || |   :::|: :  : | :: :|  :    | ||| |   |
Db     468 IKIKNEDLTFIAEKNSFSEEPFQDEIVSYNTKNKPLNFNYSLDKIIVDY---------- 516

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKK----------------YELDKYTMFHYLRAQ 103
             ||| |      :||    |  |                    : :|  |:: || ||
Db     517 ----NLQSKI-----TLPNDRTTPVTKGIPYAPEYKSNAASTIEIHNIDDNTIYQYLYAQ 567

Qy     104 EFEHGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDE 163
           :     || :||||::||:| :::|::| |  : |||:  :  :|| ||  :: |||:|
Db     568 KSPTTLQRITMTNSVDDALINSTKIYSYFPS-VISKVNQGAQGILFLQWVRDIIDDFTNE 626

Qy     164 TSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           :|: :| |||:|:: |:|||||||||    |: :|:|||  :|
Db     627 SSQKTTIDKISDVSTIVPYIGPALNIVKQGYEGNFIGALETTG 669


RESULT 11
S70582
botulinum neurotoxin type Dsa precursor - Clostridium botulinum phage d-sA
C;Species: Clostridium botulinum phage d-sA
A;Note: host Clostridium botulinum type D (strain South Africa)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C;Accession: S70582
R;Moriishi, K.; Koura, M.; Abe, N.; Fujii, N.; Fujinaga, Y.; Inoue, K.; Ogumad, K.
Biochim. Biophys. Acta 1307, 123-126, 1996
A;Title: Mosaic structures of neurotoxins produced from Clostridium botulinum types C an
A;Reference number: S70582; MUID:96283801; PMID:8679691
A;Accession: S70582
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1285 <MOR>
A;Cross-references: EMBL:D38442; NID:g1374775; PIDN:BAA07477.1; PID:g1374776
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
a disulfide bond. The heavy chain mediates the binding of toxin to the presynaptic membr
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; neurotoxin; transmembrane protein
F;1-447/Product: botulinum neurotoxin type Dsa light chain #status predicted <MAT1>
F;448-1285/Product: botulinum neurotoxin type Dsa heavy chain #status predicted <MAT2>

  Query Match             28.3%;  Score 303.5;  DB 2;  Length 1285;
  Best Local Similarity   33.2%;  Pred. No. 4e-16;
  Matches   72;  Conservative   38;  Mismatches   80;  Indels   27;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           |:| |  | :   :|: : :: : : || :||:|   :| |||   :  |      ||
Db     451 IQVKNNTLPYVADKDSISQEIFESQIITDETNVENYSDNFSLD---ESILDAKVPTNPEA 507

Qy      61 ISIENLSSDIIGQLELMPNI-----------ERFPNGKKYELDKYTMFHYLRAQEFEHGK 109
           :             |:||:           | | :    ::|    ::|| ||:  :
Db     508 VD------------PLLPNVNMEPLNVPGEEEVFYDDITKDVDYLNSYYYLEAQKLSNNV 555

Qy     110 SRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVST 169
           | || || |||   :::|||  |   :|||  :| :|| |  ::| |||    :  |
Db     556 ENITLTTSVEEALGYSNKIYTFLPS-LAEKVNKGVQAGLFLNWANEVVEDFTTNIMKKDT 614

Qy     170 TDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           |||:|:: ||||||||||||   : :|  |   :|
Db     615 LDKISDVSAIIPYIGPALNIGNSALRGNFKQAFATAG 651


RESULT 12
S11455
botulinum neurotoxin type D - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S11455
R;Binz, T.; Kurazono, H.; Popoff, M.R.; Eklund, M.W.; Sakaguchi, G.; Kozaki, S.; Krie
Nucleic Acids Res. 18, 5556, 1990
A;Title: Nucleotide sequence of the gene encoding Clostridium botulinum neurotox[redacted]ly
A;Reference number: S11455; MUID:91016853; PMID:2216736
A;Accession: S11455
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-1276 <BIN>
A;Cross-references: EMBL:X54254; NID:g40395; PIDN:CAA38175.1; PID:g40396
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             27.1%;  Score 290.5;  DB 2;  Length 1276;
  Best Local Similarity   32.6%;  Pred. No. 4.3e-15;
  Matches   71;  Conservative   39;  Mismatches   79;  Indels   29;  Gaps    5;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||| |  | :   :|: : :: : : || :||::   :  |||   :  |       ||
Db     451 IKVKNNRLPYVADKDSISQEIFENKIITDETNVQNYSDKFSLD---ESILDGQVPINPEI 507

Qy      61 ISIENLSSDIIGQLELMPNIERFP-----------NGKKYELDKYTMFHYLRAQEFEHG 108
           :              |:||:  |            :  || :|    ::|| :|:  :
Db     508 VD------------PLLPNVNMEPLNLPGEEIVFYDDITKY-VDYLNSYYYLESQKLSNN 554

Qy     109 KSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVS 168
           | || || |||   :::|||  |   :|||  :| :|| |  ::| |||    :
Db     555 VENITLTTSVEEALGYSNKIYTFLPS-LAEKVNKGVQAGLFLNWANEVVEDFTTNIMKKD 613

Qy     169 TTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           | |||:|:::||||||||||||   : :|  |   :|
Db     614 TLDKISDVSVIIPYIGPALNIGNSALRGNFNQAFATAG 651


RESULT 13
S46431
botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)
N;Alternate names: BoNT/C1 protein
C;Species: Clostridium botulinum phage 1C
A;Variety: strain C 468
C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
C;Accession: S46431; S49107
R;Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.
Mol. Gen. Genet. 243, 631-640, 1994
A;Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxi
A;Reference number: S46426; MUID:94301293; PMID:8028579
A;Accession: S46431
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1291 <HAU>
A;Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175
A;Experimental source: strain C 468
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C;Superfamily: tetanus toxin

  Query Match             24.6%;  Score 263;  DB 2;  Length 1291;
  Best Local Similarity   31.2%;  Pred. No. 7e-13;
  Matches   64;  Conservative   39;  Mismatches   94;  Indels    8;  Gaps    4;
```

```
Qy      3 VNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPENIS 62
          | | || |     :   |:   ::|  :| :    :|:|:|   |  |:    |  |:
Db    458 VKNTDLPFIGDISDVKTDIFLRKDINEETEVIYYPDNVSVD---QVILS---KNTSEHGQ 511

Qy     63 IENLSSDIIGQLELMPNIER-FPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEA 121
          :: |   |  : |::|   : | : :   :|    ::|| :|:          | |: ||
Db    512 LDLLYPSIDSESEILPGENQVFYDNRTQNVDYLNSYYYLESQKLSDNVEDFTFTRSIEEA 571

Qy    122 LLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIP 181
          | | ::|||:| :    |||   :  :|| |   :| |||       | |||:|:: |||
Db    572 LDNSAKVYTYFPT-LANKVNAGVQGGLFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP 630

Qy    182 YIGPALNIGNMLYKDDFVGALIFSG 206
          |||||||| | : : :|  |   :|
Db    631 YIGPALNISNSVRRGNFTEAFAVTG 655


RESULT 14
A49777
botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain c-s
C;Species: Clostridium botulinum phage
C;Date: 10-Mar-1994 #sequence_revision 07-Apr-1994 #text_change 23-Mar-2001
C;Accession: S11291; A35396; S22166; A49777
R;Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet, P.;
Nucleic Acids Res. 18, 4924, 1990
A;Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
A;Reference number: S11291; MUID:90370487; PMID:2204031
A;Accession: S11291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84,'P',86-1291 <HAU>
A;Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A;Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxi
A;Reference number: A35396; MUID:91024998; PMID:2222445
A;Accession: A35396
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-669,'R',671-1291 <TS1>
R;Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
submitted to the EMBL Data Library, December 1991
A;Description: Nucleotide sequence of the gene for one of the components of hemagglutini
A;Reference number: S22163
A;Accession: S22166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <TS2>
A;Cross-references: EMBL:X62389; NID:g558175; PIDN:CAA44263.1; PID:g40390
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.
Appl. Environ. Microbiol. 57, 1168-1172, 1991
A;Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and whc
A;Reference number: A49777; MUID:91282468; PMID:2059039
A;Accession: A49777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-607 <TS3>
A;Cross-references: GB:D90210
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             24.6%;  Score 263;  DB 2;  Length 1291;
  Best Local Similarity   31.2%;  Pred. No. 7e-13;
  Matches   64;  Conservative   39;  Mismatches   94;  Indels    8;  Gaps    4;

Qy      3 VNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPENIS 62
          | | || |     :   |:   ::|  :| :    :|:|:|   |  |:    |  |:
Db    458 VKNTDLPFIGDISDVKTDIFLRKDINEETEVIYYPDNVSVD---QVILS---KNTSEHGQ 511

Qy     63 IENLSSDIIGQLELMPNIER-FPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEA 121
          :: |   |  : |::|   : | : :   :|    ::|| :|:          | |: ||
Db    512 LDLLYPSIDSESEILPGENQVFYDNRTQNVDYLNSYYYLESQKLSDNVEDFTFTRSIEEA 571

Qy    122 LLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIP 181
          | | ::|||:| :    |||   :  :|| |   :| |||       | |||:|:: |||
Db    572 LDNSAKVYTYFPT-LANKVNAGVQGGLFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP 630

Qy    182 YIGPALNIGNMLYKDDFVGALIFSG 206
          |||||||| | : : :|  |   :|
Db    631 YIGPALNISNSVRRGNFTEAFAVTG 655


RESULT 15
A84977
hypothetical protein [imported] - Buchnera sp. (strain APS)
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: A84977
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: A84977
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <STO>
A;Cross-references: GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: yfiO; BU402

  Query Match             10.0%;  Score 107;  DB 2;  Length 246;
  Best Local Similarity   21.8%;  Pred. No. 0.24;
  Matches   44;  Conservative   40;  Mismatches   74;  Indels   44;  Gaps    8;

Qy     14 EDNFTNDLNKGEEITSDTN-IEAAEENISLDLIQQYYLTFNFDNEPENIS---------- 62
          ::|| | ::  |:|  : |    : : | :|||  ||   |||   :||
Db     46 KENFDNAISILEKIKKNNNTANISNDKIQIDLIYAYYKILNFDQARKNIEEFMYFYPNHP 105

Qy     63 ----------IENLSSDIIGQLELMPNIERFPNGKKYELD-----KYTMFHYLRAQEFEH 107
                     : ::| |      : | |  : |   |  :     || :: | :::   :
Db    106 NIDYVVYIQCLISMSLDKNRFFSVFP-INYYKNDYFYAKNAFFQLKYFIYQYPKSRYVVN 164

Qy    108 GKSR-IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATE----------AAMFLGWVEQ- 155
           |   | : | ::|  |:  : | ||  :|:  :|:  |          |   | ::|:
Db    165 AKKNLIYIKNRLSEHDLSILKFY-FFHKEYIAVINRGEEMLQRYSETPSARKALIYIEKS 223

Qy    156 ----LVYDFTDETSEVSTTDKI 173
               ::|   : |::   :||
Db    224 YYALKIFDTAKKISKIILLNKI 245


Search completed: March 13, 2003, 11:41:44
Job time : 21.1066 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:47 ; Search time 31.0202 Seconds
(without alignments)
1374.968 Million cell updates/sec

Title: US-09-917-791-21
Perfect score: 1071
Sequence: 1 IKVNNWDLFFSPSEDNFTND.....NIGNMLYKDDFVGALIFSGA 207

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	407.5	38.0	1268	2	Q45851	Q45851 clostridium
2	374.5	35.0	1255	2	Q9FAR6	Q9far6 clostridium
3	368.5	34.4	1251	2	Q9K395	Q9k395 clostridium
4	367.5	34.3	1291	2	Q933K0	Q933k0 clostridium
5	366.5	34.2	1291	2	Q93G71	Q93g71 clostridium
6	366.5	34.2	1291	2	Q9ZAJ8	Q9zaj8 clostridium
7	364.5	34.0	1278	2	Q57236	Q57236 clostridium
8	357.5	33.4	1291	2	Q08077	Q08077 clostridium
9	347.5	32.4	1280	2	Q9ZAJ5	Q9zaj5 clostridium
10	333	31.1	173	2	Q9KW88	Q9kw88 clostridium
11	331	30.9	173	2	Q9KW90	Q9kw90 clostridium
12	330	30.8	173	2	Q9K345	Q9k345 clostridium
13	330	30.8	173	2	Q9KW89	Q9kw89 clostridium
14	314	29.3	1310	2	Q93N27	Q93n27 clostridium
15	303.5	28.3	1285	2	Q9LBR1	Q9lbr1 clostridium
16	303.5	28.3	1285	2	Q45967	Q45967 clostridium
17	290.5	27.1	1275	12	Q9QTG7	Q9qtg7 clostridium
18	263	24.6	1280	2	Q45849	Q45849 clostridium
19	263	24.6	1291	2	Q93HT3	Q93ht3 clostridium
20	262	24.5	1280	2	Q9LBS7	Q9lbs7 clostridium
21	106.5	9.9	1197	2	Q45888	Q45888 clostridium
22	106	9.9	1196	2	Q53550	Q53550 clostridium
23	106	9.9	1196	2	Q45916	Q45916 clostridium
24	106	9.9	1196	2	Q93HT4	Q93ht4 clostridium
25	106	9.9	1196	9	Q9ZX77	Q9zx77 clostridium
26	106	9.9	1196	9	Q38197	Q38197 clostridium
27	105	9.8	440	5	Q9GSC4	Q9gsc4 plasmodium
28	105	9.8	440	5	Q9GSC0	Q9gsc0 plasmodium
29	105	9.8	440	5	Q9GSB8	Q9gsb8 plasmodium
30	105	9.8	2771	5	Q26216	Q26216 plasmodium
31	104	9.7	1193	2	P71108	P71108 clostridium
32	103	9.6	432	5	Q9GSC5	Q9gsc5 plasmodium
33	103	9.6	432	5	Q9GSC1	Q9gsc1 plasmodium
34	103	9.6	432	5	Q9GSB9	Q9gsb9 plasmodium
35	103	9.6	1163	2	Q45850	Q45850 clostridium
36	101	9.4	432	5	Q9GSC8	Q9gsc8 plasmodium
37	101	9.4	440	5	Q9GSC7	Q9gsc7 plasmodium
38	101	9.4	1198	2	O06018	O06018 clostridium
39	100.5	9.4	356	2	Q9AKM7	Q9akm7 rickettsia
40	100	9.3	664	5	O97263	O97263 plasmodium
41	99.5	9.3	1055	12	Q65146	Q65146 african swi
42	99.5	9.3	2747	5	Q9BJX9	Q9bjx9 plasmodium
43	99	9.2	1193	2	Q45914	Q45914 clostridium
44	99	9.2	1193	2	P71107	P71107 clostridium
45	99	9.2	1196	2	Q9LBS8	Q9lbs8 clostridium

ALIGNMENTS

RESULT 1
Q45851
ID Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Neurotoxin type F.
GN BONT /F.
OS Clostridium baratii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93252228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium barati type F
RT neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEMS Microbiol. Lett. 108:175-182(1993).
DR EMBL; X68262; CAA48329.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 38.0%; Score 407.5; DB 2; Length 1268;
Best Local Similarity 44.0%; Pred. No. 3.1e-22;
Matches 95; Conservative 39; Mismatches 61; Indels 21; Gaps 7;

Qy 1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
Db 436 IKVNNRDLFFVASESSYNENGINSPKEI-DDTTITNNNYKKNLD---EVILDYNSD---- 487

```
Qy      60 NISIENLSSDIIGQL----ELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           :| |||| ::         :|  :   ||    |:| :|| :| || ||:   |:|
Db     488 --AIPNLSSRLLNTTAQNDSYVPKYD--SNGTSEIKEYTVDKLNVFFYLYAQKAPEGESA 543

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           |:||:||| |||: |:||||||||::  |||  :||:|: |::|:: ||| | :: || |
Db     544 ISLTSSVNTALLDASKVYTFFSSDFINTVNKPVQAALFISWIQQVINDFTTEATQKSTID 603

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||::|:||:| |||||| : | :|  |:   ||
Db     604 KIADISLIVPYVGLALNIGNEVQKGNFKEAIELLGA 639


RESULT 2
Q9FAR6
ID   Q9FAR6      PRELIMINARY;      PRT;  1255 AA.
AC   Q9FAR6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Type E botulinum toxin.
GN   BONT/E.
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BL 6340/ATCC 43755/BL 5520/KZ 147;
RX   MEDLINE=20509829; PubMed=11055954;
RA   Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y.,
RA   Yamakawa K., Oguma K., Sakaguchi Y., Nakamura S.;
RT   "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium
RT   butyricum Strains.";
RL   Appl. Environ. Microbiol. 66:4992-4997(2000).
DR   EMBL; AB039264; BAB12249.1; -.
DR   HSSP; P10845; 3BTA.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1255 AA;  143918 MW;  1B557B9D85CD8E4D CRC64;

  Query Match             35.0%;  Score 374.5;  DB 2;  Length 1255;
  Best Local Similarity   38.9%;  Pred. No. 9.2e-20;
  Matches   84;  Conservative   51;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
           |::|| :|||  ||::: :| :|  :||    ||: | |        : : |  | || :
Db     430 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 481

Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           : |  :| | |:  |      :|  :   ||    :::::::  :| || ||:   |::
Db     482 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 537

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           : ||:|:: |||   ::||||||:::  |||  :||:|:||::|:: ||| | :: || |
Db     538 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQKSTVD 597

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||:|::|||| ||||||   | :|  ||   ||
Db     598 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 633


RESULT 3
Q9K395
ID   Q9K395      PRELIMINARY;      PRT;  1251 AA.
AC   Q9K395;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Type E botulinum toxin.
GN   BONT/E.
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 095;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (LCL 095) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 155;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
RA   Kato H., Nakamura S., Karasawa T.;
RT   "C. butyricum (LCL 155) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1899;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1897;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1898;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1886;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1886) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1887;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1889;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1889) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1890;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1890) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1891;
```

```
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1891) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 063;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (LCL 063) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037714; BAB03522.1; -.
DR   EMBL; AB037704; BAB03512.1; -.
DR   EMBL; AB037705; BAB03513.1; -.
DR   EMBL; AB037706; BAB03514.1; -.
DR   EMBL; AB037707; BAB03515.1; -.
DR   EMBL; AB037708; BAB03516.1; -.
DR   EMBL; AB037709; BAB03517.1; -.
DR   EMBL; AB037710; BAB03518.1; -.
DR   EMBL; AB037711; BAB03519.1; -.
DR   EMBL; AB037712; BAB03520.1; -.
DR   EMBL; AB037713; BAB03521.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1251 AA;  143751 MW;  2021F4E427070296 CRC64;

  Query Match             34.4%;  Score 368.5;  DB 2;  Length 1251;
  Best Local Similarity   38.4%;  Pred. No. 2.6e-19;
  Matches   83;  Conservative   51;  Mismatches   61;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
           |::|| :|||  ||::: :| :|  :||    ||: | |        : : |  | || :
Db     427 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 478

Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           : |  :| | |:  |      :|  :   ||    :::::::  :| || ||:   |::
Db     479 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 534

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           : ||:|:: |||   ::|||||::::  |||  :||:|: |::|:: ||| | :: || |
Db     535 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVD 594

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||:|::|||| ||||||   | :|  ||   ||
Db     595 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 630


RESULT 4
Q933K0
ID   Q933K0      PRELIMINARY;      PRT;  1291 AA.
AC   Q933K0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Type B cryptic neurotoxin.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=593, AND 588;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF300466; AAL11499.1; -.
DR   EMBL; AF300465; AAL11498.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin.
SQ   SEQUENCE   1291 AA;  150843 MW;  7AC1737B0FA5A151 CRC64;

  Query Match             34.3%;  Score 367.5;  DB 2;  Length 1291;
  Best Local Similarity   38.3%;  Pred. No. 3.2e-19;
  Matches   85;  Conservative   40;  Mismatches   62;  Indels   35;  Gaps    3;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISL-DLIQQYYLTFNFDNEPE 59
           | |:| ||||   :::|::||:| | |  :|     | : |: :||
Db     447 IDVDNEDLFFIADKNSFSDDLSKNERIAYNTQNNYIENDFSINELI-------------- 492

Qy      60 NISIENLSSDIIGQLEL-------------MPNIERFPNGKKYELDKYTMFHYLRAQEF 105
             | :|:| ::||                :|  :: |  ||   |: |:| || :| |
Db     493 ------LDTDLISKIELPSENTESLTDFNVYVPVYKKQPAIKKIFTDENTIFQYLYSQTF 546

Qy     106 EHGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETS 165
             |:||:| ::|||  ::||:||| ||:|  ||  || :| |||:|:| ||  | :
Db     547 PLDIRDISLTSSFDDALLFSNKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVNDFVIEAN 606

Qy     166 EVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           : || ||||||::|:|||| |||:||   | :|  |   :||
Db     607 KSSTMDKIADISLIVPYIGLALNVGNETAKGNFENAFEIAGA 648


RESULT 5
Q93G71
ID   Q93G71      PRELIMINARY;      PRT;  1291 AA.
AC   Q93G71;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Neurotoxin type B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1436;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF295926; AAK97132.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1291 AA;  150824 MW;  D7CA07BAE2EB8CD2 CRC64;

  Query Match             34.2%;  Score 366.5;  DB 2;  Length 1291;
  Best Local Similarity   38.5%;  Pred. No. 3.8e-19;
  Matches   85;  Conservative   41;  Mismatches   62;  Indels   33;  Gaps    3;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           | |:| ||||   :::|::||:| | |  :|          |  |:  :|     :
Db     447 IDVDNEDLFFIADKNSFSDDLSKNERIAYNT-------------QNNYIDNDF-----S 487

Qy      61 ISIENLSSDIIGQLEL-------------MPNIERFPNGKKYELDKYTMFHYLRAQEFE 106
           |:   | :|:| ::||               :|  :: |  ||   |: |:| || :| |
Db     488 INELILDTDLISKIELPSENTESLTDFNVYVPEYKKQPAIKKIFTDENTIFQYLYSQTFP 547

Qy     107 HGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSE 166
             |:||:| ::|||  ::||:||| ||:|  ||  || :| |||:|:| ||  | ::
Db     548 LDIRDISLTSSFDDALLFSNKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVDDFVIEANK 607
```

```
Qy    167 VSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
          || |||||||::|:|||| |||:||   | :|  |   :||
Db    608 SSTMDKIADISLIVPYIGLALNVGNETAKGNFENAFEIAGA 648


RESULT 6
Q9ZAJ8
ID   Q9ZAJ8      PRELIMINARY;      PRT;  1291 AA.
AC   Q9ZAJ8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13630; CAA73968.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1291 AA;  150840 MW;  E4D3B0E46AB2E735 CRC64;

  Query Match           34.2%;  Score 366.5;  DB 2;  Length 1291;
  Best Local Similarity  38.3%;  Pred. No. 3.8e-19;
  Matches   85;  Conservative   40;  Mismatches   62;  Indels   35;  Gaps    3;

Qy      1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISL-DLIQQYYLTFNFDNEPE 59
          | |:| ||||   :::|::||:| | |  :|     | : |: :||
Db    447 IDVDNEDLFFIADKNSFSDDLSKNERIAYNTQNNYIENDFSINELI-------------- 492

Qy     60 NISIENLSSDIIGQLEL--------------MPNIERFPNGKKYELDKYTMFHYLRAQEF 105
                | :|:| ::||              :|  :: |  ||   |: |:| || :| |
Db    493 ------LDTDLISKIELPSENTESLTDFNVYVPVYKKQPAIKKIFTDENTIFQYLYSQTF 546

Qy    106 EHGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETS 165
              |:||:| ::|||  ::||:||| ||:|  ||  || :| |||:|:| ||  | :
Db    547 PLDIRDISLTSSFDDALLFSNKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVDDFVIEAN 606

Qy    166 EVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
          : || |||||||::|:|||| |||:||   | :|  |   :||
Db    607 KSSTMDKIADISLIVPYIGLALNVGNETAKGNFENAFEIAGA 648


RESULT 7
Q57236
ID   Q57236      PRELIMINARY;      PRT;  1278 AA.
AC   Q57236; Q45863;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Botulinum neurotoxin type F (BONT/F protein).
GN   BONT/F.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 10281;
RA   Hutson R.A., Collins M.D.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 635-1000 FROM N.A.
RC   STRAIN=NCTC 1028;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-27 FROM N.A.
RC   STRAIN=LANGELAND;
RX   MEDLINE=98404102; PubMed=9732534;
RA   East A.K., Bhandari M., Hielm S., Collins M.D.;
RT   "Analysis of the botulinum neurotoxin type F gene clusters in
RT   proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT   barati.";
RL   Curr. Microbiol. 37:262-268(1998).
DR   EMBL; X81714; CAA57358.1; -.
DR   EMBL; L35496; AAA23210.1; -.
DR   EMBL; X70821; CAA50152.1; -.
DR   EMBL; X99064; CAA67512.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin.
SQ   SEQUENCE   1278 AA;  147073 MW;  A1BE1318431D6918 CRC64;

  Query Match           34.0%;  Score 364.5;  DB 2;  Length 1278;
  Best Local Similarity  38.0%;  Pred. No. 5.3e-19;
  Matches   81;  Conservative   52;  Mismatches   65;  Indels   15;  Gaps    7;

Qy      1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
          |:||| :|||  || ::  ||:|  :||   ||:     | :||   :  | :| :  |:
Db    446 IRVNNRELFFVASESSYNENDINTPKEIDDTTNLNNNYRN-NLD---EVILDYNSETIPQ 501

Qy     60 NISIENLSSDIIGQLELMPNIERFPNGKKYELDKY-----TMFHYLRAQEFEHGKSRIAL 114
          || : |:: ::     :|   |: :    |::::      :| || ||:   |:: |:|
Db    502 -ISNQTLNT-LVQDDSYVP---RYDSNGTSEIEEHNVVDLNVFFYLHAQKVPEGETNISL 556

Qy    115 TNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA 174
          |:|:: ||   |:|||||||:::  :||   ||:|: |: |:: ||| | :: || ||||
Db    557 TSSIDTALSEESQVYTFFSSEFINTINKPVHAALFISWINQVIRDFTTEATQKSTFDKIA 616

Qy    175 DITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
          ||::::||:| |||||| : |::|  |    ||
Db    617 DISLVVPYVGLALNIGNEVQKENFKEAFELLGA 649


RESULT 8
Q08077
ID   Q08077      PRELIMINARY;      PRT;  1291 AA.
AC   Q08077;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Botulinum neurotoxin type B (EC 3.4.24.-) (BONT/B).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
```

```
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EKLUND 17B ATCC25765;
RX   MEDLINE=94122659; PubMed=7764370;
RA   Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT   "Nucleotide sequence of the gene coding for non-proteolytic
RT   Clostridium botulinum type B neurotoxin: comparison with other
RT   clostridial neurotoxins.";
RL   Curr. Microbiol. 28:101-110(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES SYNAPTOBREVIN-2.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC       NEUROTOXIN.
CC   -!- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC       REGION.
DR   EMBL; X71343; CAA50482.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
SQ   SEQUENCE   1291 AA;  150513 MW;  71BCAFE23D69FAAA CRC64;

  Query Match             33.4%;  Score 357.5;  DB 2;  Length 1291;
  Best Local Similarity   38.5%;  Pred. No. 1.8e-18;
  Matches   85;  Conservative   41;  Mismatches   62;  Indels   33;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           | |:| :|||   :::|::||:| | :  :|              |  |:  :|   | |
Db     447 IDVDNENLFFIADKNSFSDDLSKNERVEYNT-------------QNNYIGNDF---PIN 489

Qy      61 ISIENLSSDIIGQLEL-------------MPNIERFPNGKKYELDKYTMFHYLRAQEFE 106
             |  | :|:| ::||              :|  |: |  ||   |: |:| || :| |
Db     490 ELI--LDTDLISKIELPSENTESLTDFNVDVPVYEKQPAIKKVFTDENTIFQYLYSQTFP 547

Qy     107 HGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSE 166
               |:||:| ::|||  |:||:||| ||:|  ||  || :| |||:|:| ||  | ::
Db     548 LNIRDISLTSSFDDALLVSSKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVDDFVIEANK 607

Qy     167 VSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           || ||||||::|:|||| |||:|:   | :|  |   :|:
Db     608 SSTMDKIADISLIVPYIGLALNVGDETAKGNFESAFEIAGS 648


RESULT 9
Q9ZAJ5
ID   Q9ZAJ5      PRELIMINARY;      PRT;  1280 AA.
AC   Q9ZAJ5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13631; CAA73972.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1280 AA;  147487 MW;  D0F748976EBC222C CRC64;

  Query Match             32.4%;  Score 347.5;  DB 2;  Length 1280;
  Best Local Similarity   37.7%;  Pred. No. 1e-17;
  Matches   80;  Conservative   50;  Mismatches   69;  Indels   13;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
           |:||| :||| || ::  :|:|  :||   ||:     | :||   :  | :| :  |:
Db     446 IRVNNRELFFVASESSYNESDINTPKEIDDTTNLNNNYRN-NLD---EVILDYNSETIPQ 501

Qy      60 NISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSRIALT 115
             ||   |:: ::     :|  :   ||    ::|::  : :| || ||:   |:: |:||
Db     502 -ISNRTLNT-LVQDNSYVPRYD--SNGTSEIEEYDVVDFNVFFYLHAQKVPEGETNISLT 557

Qy     116 NSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIAD 175
           :|:: |||  |:|||||||:::  :||   ||:|: |: ::: ||| | :: || ||||
Db     558 SSIDTALLEESKVYTFFSSEFIDTINKPVNAALFIDWISKVIRDFTTEATQKSTVDKIAD 617

Qy     176 ITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |::|:||:| ||||     | :|  |    ||
Db     618 ISLIVPYVGLALNIVIEAEKGNFEEAFELLGA 649


RESULT 10
Q9KW88
ID   Q9KW88      PRELIMINARY;      PRT;  173 AA.
AC   Q9KW88;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Botulinum neurotoxin type E (Fragment).
GN   BONT/E.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=164-1;
RX   MEDLINE=20575211; PubMed=11133447;
RA   Kimura B., Kawasaki S., Nakano H., Fujii T.;
RT   "Rapid, Quantitative PCR Monitoring of Growth of Clostridium botulinum
RT   Type E in Modified-Atmosphere-Packaged Fish.";
RL   Appl. Environ. Microbiol. 67:206-216(2001).
DR   EMBL; AB040128; BAB07890.2; -.
DR   HSSP; P10845; 3BTA.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19121 MW;  42F5822DE3B01F1F CRC64;

  Query Match             31.1%;  Score 333;  DB 2;  Length 173;
  Best Local Similarity   35.7%;  Pred. No. 1.1e-17;
  Matches   71;  Conservative   49;  Mismatches   47;  Indels   32;  Gaps    5;
```

```
Qy      13 SEDNFTNDLNKGEEITSDTNIEAA----EENISLDLIQQYYLTFNFDNEPENISIENLSS 68
           | :|: |||   :::  : | |:|    :: ::| :    |:   :|:        | :|
Db       1 SNNNYENDL---DQVILNFNSESAPGLSDKKLNLTIQDDAYIP-KYDS--------NGTS 48

Qy      69 DIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEALLNPSRV 128
           ||                :::::::  :| || ||:   |:: : ||:|:: |||   ::
Db      49 DI----------------EQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTALLEQPKI 92

Qy     129 YTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALN 188
           ||||||::  |||  :||:|:||::|:: ||| | :: || |||||:|::|||| |||
Db      93 YTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQKSTVDKIADISIVVPYIGLALN 152

Qy     189 IGNMLYKDDFVGALIFSGA 207
           |||   | :|  ||   ||
Db     153 IGNEAQKGNFKDALELLGA 171


RESULT 11
Q9KW90
ID   Q9KW90      PRELIMINARY;      PRT;   173 AA.
AC   Q9KW90;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Botulinum neurotoxin type E (Fragment).
GN   BONT/E.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BIWAKO;
RX   MEDLINE=20575211; PubMed=11133447;
RA   Kimura B., Kawasaki S., Nakano H., Fujii T.;
RT   "Rapid, Quantitative PCR Monitoring of Growth of Clostridium botulinum
RT   Type E in Modified-Atmosphere-Packaged Fish.";
RL   Appl. Environ. Microbiol. 67:206-216(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=35396;
RA   Kimura B.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=35396;
RA   Kmura B., Kawasaki S., Nakano H., Fujii T.;
RT   "Real-time, quantitative PCR assay of Clostriium botulinum type E in
RT   modified-atmosphere-packaged fish.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB040125; BAB07887.2; -.
DR   EMBL; AB040127; BAB07889.2; -.
DR   HSSP; P10845; 3BTA.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19151 MW;  41781916CF06AB1F CRC64;

  Query Match             30.9%;  Score 331;  DB 2;  Length 173;
  Best Local Similarity   35.7%;  Pred. No. 1.5e-17;
  Matches   71;  Conservative   48;  Mismatches   48;  Indels   32;  Gaps    5;

Qy      13 SEDNFTNDLNKGEEITSDTNIEAA----EENISLDLIQQYYLTFNFDNEPENISIENLSS 68
           | :|: |||   :::  : | |:|    :| ::| :    |:   :|:        | :|
Db       1 SNNNYENDL---DQVILNFNSESAPGLSDEKLNLTIQNDAYIP-KYDS--------NGTS 48

Qy      69 DIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEALLNPSRV 128
           ||                :::::::  :| || ||:   |:: : ||:|:: |||   ::
Db      49 DI----------------EQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTALLEQPKI 92

Qy     129 YTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALN 188
           ||||||::  |||  :||:|: |::|:: ||| | :: || |||||:|::|||| |||
Db      93 YTFFSSEFINNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVDKIADISIVVPYIGLALN 152

Qy     189 IGNMLYKDDFVGALIFSGA 207
           |||   | :|  ||   ||
Db     153 IGNEAQKGNFKDALELLGA 171


RESULT 12
Q9K345
ID   Q9K345      PRELIMINARY;      PRT;   173 AA.
AC   Q9K345; Q9KW91;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Botulinum neurotoxin type E (Fragment).
GN   BONT/E.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IWANAI, AND TENNO2;
RX   MEDLINE=20575211; PubMed=11133447;
RA   Kimura B., Kawasaki S., Nakano H., Fujii T.;
RT   "Rapid, Quantitative PCR Monitoring of Growth of Clostridium botulinum
RT   Type E in Modified-Atmosphere-Packaged Fish.";
RL   Appl. Environ. Microbiol. 67:206-216(2001).
DR   EMBL; AB040123; BAB07885.2; -.
DR   EMBL; AB040124; BAB07886.2; -.
DR   HSSP; P10845; 3BTA.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19061 MW;  5A89702DE3ACED16 CRC64;

  Query Match             30.8%;  Score 330;  DB 2;  Length 173;
  Best Local Similarity   34.1%;  Pred. No. 1.8e-17;
  Matches   72;  Conservative   42;  Mismatches   41;  Indels   56;  Gaps    4;

Qy      13 SEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE------NISIE-- 64
           | :|: |||:                          |  | || :: |       |::|:
Db       1 SNNNYENDLD----------------------QVILNFNSESAPGLSDKKLNLTIQDD 36

Qy      65 --------NLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTN 116
                  | :|||                 :::::::  :| || ||:   |:: : ||:
Db      37 ACIPKYDSNGTSDI----------------EQHDVNELNVFFYLDAQKVPEGENNVNLTS 80

Qy     117 SVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADI 176
           |:: |||   ::||||||:::  |||  :||:|:||::|:: ||| | :: || ||||||
Db      81 SIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQKSTVDKIADI 140

Qy     177 TIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           :|::|||| ||||||   | :|  ||   ||
Db     141 SIVVPYIGLALNIGNEAQKGNFKDALELLGA 171


RESULT 13
Q9KW89
ID   Q9KW89      PRELIMINARY;      PRT;   173 AA.
AC   Q9KW89;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Botulinum neurotoxin type E (Fragment).
GN   BONT/E.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=5545;
RX   MEDLINE=20575211; PubMed=11133447;
```

```
RA   Kimura B., Kawasaki S., Nakano H., Fujii T.;
RT   "Rapid, Quantitative PCR Monitoring of Growth of Clostridium botulinum
RT   Type E in Modified-Atmosphere-Packaged Fish.";
RL   Appl. Environ. Microbiol. 67:206-216(2001).
DR   EMBL; AB040126; BAB07888.2; -.
DR   HSSP; P10845; 3BTA.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19137 MW;  4178024F5BC4BF1F CRC64;

  Query Match             30.8%;  Score 330;  DB 2;  Length 173;
  Best Local Similarity   35.7%;  Pred. No. 1.8e-17;
  Matches   71;  Conservative   48;  Mismatches   48;  Indels   32;  Gaps    5;

Qy      13 SEDNFTNDLNKGEEITSDTNIEAA----EENISLDLIQQYYLTFNFDNEPENISIENLSS 68
           | :|: |||   :::  : | |:|    :| ::| :    |:   :|:        | :|
Db       1 SNNNYENDL---DQVILNFNSESAPGLSDEKLNLTIQNDAYIP-KYDS--------NGTS 48

Qy      69 DIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEALLNPSRV 128
           ||                 :::::::  :| || ||:   |:: : ||:|:: |||   ::
Db      49 DI----------------EQHDVNELNVFFYLDAQKVPEGENNVNLTSSIDTALLEQPKI 92

Qy     129 YTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALN 188
           ||||||:::  ||| :||:|:||::|:: ||| | :: || ||||||:|::|| | |||
Db      93 YTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQESTVDKIADISIVVPYKGLALN 152

Qy     189 IGNMLYKDDFVGALIFSGA 207
           |||   | :|  ||   ||
Db     153 IGNEAQKGNFKDALELLGA 171


RESULT 14
Q93N27
ID   Q93N27      PRELIMINARY;      PRT;  1310 AA.
AC   Q93N27;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Tetanus toxin (Fragment).
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shumin Z., Dianliang L.;
RT   "Cloning and sequence analysis of tetanus toxin gene.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389424; AAK72964.2; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER    1310   1310
SQ   SEQUENCE   1310 AA;  150316 MW;  9EADDC914418E450 CRC64;

  Query Match             29.3%;  Score 314;  DB 2;  Length 1310;
  Best Local Similarity   33.2%;  Pred. No. 3.4e-15;
  Matches   74;  Conservative   44;  Mismatches   67;  Indels   38;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||: | || |   :::|: :  : | :: :|  :    | ||| |   |
Db     469 IKIKNEDLIFIAEKNSFSEEPFQDEIVSYNTKNKPLNFNYSLDKIILDY---------- 517

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKK----------------YELDKYTMFHYLRAQ 103
              || | |      :||    |  |                  : :|  |:: || ||
Db     518 ----NLQSKI-----TLPNDRTTPVTKGIPYAPEYKSNAASTIEIHNIDDNTIYQYLYAQ 568

Qy     104 EFEHGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDE 163
           :      || :||||::||:| :::|::| |  : |||:  :  :|| ||  :: |||:|
Db     569 KSPTTLQRITMTNSVDDALINSTKIYSYFPS-VISKVNQGAQGILFLQWVRDIIDDFTNE 627

Qy     164 TSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           :|: :| |||:|:: |:|||||||||    |: :|:|||  :|
Db     628 SSQKTTIDKISDVSTIVPYIGPALNIVKQGYEGNFIGALETTG 670


RESULT 15
Q9LBR1
ID   Q9LBR1      PRELIMINARY;      PRT;  1285 AA.
AC   Q9LBR1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Neurotoxin.
GN   NT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D-4947;
RA   Sagane Y., Watanabe T., Kouguchi H., Yamamoto T., Takizawa J.,
RA   Kawabe T., Murakami F., Muroga A., Nakatsuka M., Ohyama T.;
RT   "Characterization of the Progenitor Toxin Components Produced by
RT   Clostridium botulinum Type D Strain 4947.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037920; BAA90661.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1285 AA;  147352 MW;  B63AFA487D570680 CRC64;

  Query Match             28.3%;  Score 303.5;  DB 2;  Length 1285;
  Best Local Similarity   33.2%;  Pred. No. 2e-14;
  Matches   72;  Conservative   38;  Mismatches   80;  Indels   27;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           |:| |  | :   :|: : :: : : || :||:|   :| |||   :  |       ||
Db     451 IQVKNNTLPYVADKDSISQEIFESQIITDETNVENYSDNFSLD---ESILDAKVPTNPEA 507

Qy      61 ISIENLSSDIIGQLELMPNI----------ERFPNGKKYELDKYTMFHYLRAQEFEHGK 109
           :          |:||:           | | :    ::|    ::|| ||:  :
Db     508 VD----------PLLPNVNMEPLNVPGEEEVFYDDITKDVDYLNSYYYLEAQKLSNNV 555

Qy     110 SRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVST 169
           | || || |||   :::|||  |   :|||| :| :|| |  ::| |||    :  |
Db     556 ENITLTTSVEEALGYSNKIYTFLPS-LAEKVNKGVQAGLFLNWANEVVEDFTTNIMKKDT 614

Qy     170 TDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           |||:|:: ||||||||||||   : :|  |   :|
Db     615 LDKISDVSAIIPYIGPALNIGNSALRGNFKQAFATAG 651


Search completed: March 13, 2003, 11:40:11
Job time : 35.5202 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:12 ; Search time 26.2248 Seconds
(without alignments)
711.353 Million cell updates/sec

Title: US-09-917-791-22
Perfect score: 727
Sequence: 1 LNSSLYRGTKFIIKKYASGN.....QIERSSRTLGCSWEFIPVDD 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	727	100.0	140	22	AAG79296	Amino acid sequenc
2	727	100.0	206	21	AAY77144	Botulinum neurotox
3	727	100.0	382	21	AAB36303	BoNT/A prototoxin
4	727	100.0	415	22	AAB04083	Botulism toxin C f
5	727	100.0	432	21	AAY77142	Native botulinum n
6	727	100.0	434	22	AAB04089	Botulism toxin hea
7	727	100.0	435	22	AAB04090	Botulism toxin hea
8	727	100.0	437	22	AAB04088	Botulism toxin hea
9	727	100.0	438	17	AAR95008	Type A neurotoxin
10	727	100.0	438	19	AAW68389	Clostridium botuli
11	727	100.0	438	21	AAY77134	Synthetic botulinu
12	727	100.0	445	19	AAW68391	Clostridium botuli
13	727	100.0	462	17	AAR95009	Type A neurotoxin
14	727	100.0	462	19	AAW68390	Clostridium botuli
15	727	100.0	837	21	AAY77140	Native botulinum n
16	727	100.0	1067	21	AAY93307	A manganese supero
17	727	100.0	1092	21	AAY93310	A manganese supero
18	727	100.0	1296	17	AAR95010	C. botulinum type
19	701	96.4	1295	23	AAU99339	Clostridium botuli
20	698.5	96.1	847	22	AAB04081	Botulism toxin hea
21	284.5	39.1	432	21	AAY77138	Synthetic botulinu
22	277.5	38.2	143	18	AAW09017	Immunogenic type F
23	277.5	38.2	431	18	AAW09014	Immunogenic type F
24	277.5	38.2	432	22	AAB04096	Botulism toxin hea
25	277.5	38.2	432	22	AAB04103	Botulism toxin hea
26	277.5	38.2	645	22	AAE07894	Modified clostridi
27	277.5	38.2	685	22	AAE07893	Modified clostridi
28	277.5	38.2	862	22	AAE07890	Modified clostridi
29	277.5	38.2	887	22	AAE07892	Modified clostridi
30	277.5	38.2	1032	22	AAE07901	C. botulinum C2 tr
31	277.5	38.2	1059	21	AAY93309	A manganese supero
32	277.5	38.2	1084	21	AAY93312	A manganese supero
33	277.5	38.2	1092	22	AAE07900	C. botulinum C2 tr
34	258.5	35.6	448	19	AAW68399	Clostridium botuli
35	206.5	28.4	419	22	AAB04095	Botulism toxin hea
36	206.5	28.4	449	21	AAY77137	Synthetic botulinu
37	206.5	28.4	449	22	AAB04094	Botulism toxin hea
38	206.5	28.4	452	19	AAW68396	Clostridium botuli
39	190	26.1	451	19	AAW68395	Clostridium botuli
40	169.5	23.3	440	21	AAY77135	Synthetic botulinu
41	169.5	23.3	440	22	AAB04091	Botulism toxin hea
42	169.5	23.3	472	19	AAW68393	Clostridium botuli
43	165.5	22.8	472	19	AAW68394	Clostridium botuli
44	165.5	22.8	1070	21	AAY93308	A manganese supero
45	165.5	22.8	1095	21	AAY93311	A manganese supero

ALIGNMENTS

RESULT 1
AAG79296
ID AAG79296 standard; Protein; 140 AA.
XX
AC AAG79296;
XX
DT 03-JAN-2002 (first entry)
XX
DE Amino acid sequence of botulinum neurotoxin (BoNT) fragment.
XX
KW Neurotoxin; BoNT; vaccine; botulism.
XX
OS Clostridium botulinum.
XX
PN US6287566-B1.
XX
PD 11-SEP-2001.
XX
PF 19-MAY-1995; 95US-0446114.
XX
PR 19-MAY-1995; 95US-0446114.
XX
PA (USSA) US SEC OF ARMY.
XX
PI Dertzbaugh MT;
XX
DR WPI; 2001-615462/71.
XX
PT New protective epitopes of neurotoxin of Clostridium botulinum, useful
PT for developing vaccines against neurotoxins of Clostridium botulinum
PT and for identifying protective antibodies -
XX
PS Claim 1; Column 25; 14pp; English.

```
XX
CC   The present sequence represents a fragment of the Clostridium botulinum
CC   neurotoxin (BoNT). It was produced by amplifying overlapping fragments
CC   of the BoNT gene. The amplified fragments were cloned expressed to
CC   identify immunogenic polypeptides which are capable of giving rise to
CC   protective antibodies. The BoNT polypeptide fragment are useful as
CC   vaccines, for immunizing against botulism, and as diagnostic agents
CC   to identify protective antibodies.
XX
SQ   Sequence   140 AA;

  Query Match           100.0%;  Score 727;  DB 22;  Length 140;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    121 QIERSSRTLGCSWEFIPVDD 140


RESULT 2
AAY77144
ID   AAY77144 standard; Protein; 206 AA.
XX
AC   AAY77144;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Botulinum neurotoxin serotype A (BoNTA) C-terminal subfragment AsubHc2.
XX
KW   Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW   C-terminal subfragment; Hc; botulism; VEE;
KW   Venezuelan equine encephalitis virus replicon; vaccine; diagnosis;
KW   drug screening.
XX
OS   Clostridium botulinum.
OS   Synthetic.
XX
PN   WO200002524-A2.
XX
PD   20-JAN-2000.
XX
PF   09-JUL-1999;   99WO-US15570.
XX
PR   10-JUL-1998;   98US-0092416.
PR   12-MAY-1999;   99US-0133870.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Lee JS,  Pushko P,  Smith JF,  Parker M,  Dertzbaugh MT,  Smith L;
XX
DR   WPI; 2000-160827/14.
DR   N-PSDB; AAZ87222.
XX
PT   Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT   toxin serotypes A-G, is used for inducing an immune response against
PT   botulinum -
XX
PS   Disclosure; Page 53; 54pp; English.
XX
CC   The invention relates to novel vaccines that induce a protective immune
CC   response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC   and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC   DNA construct comprising a vector, and at least one nucleic acid
CC   fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC   serotypes A-G. In preferred embodiments of the invention, the vector is a
CC   Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC   vector results in the production of large amounts of a protein encoded by
CC   a sequence cloned into the replicon. The constructs are used to produce
CC   vaccines against botulism. The proteins can also be used as diagnostic
CC   tools for the diagnosis of botulism. The transformed host cells can be
CC   used to analyse the effectiveness of drugs and agents which inhibit toxin
CC   effects. The vaccine currently used against botulism is dangerous
CC   and expensive to produce, and contains formalin, which is very painful
CC   for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC   the 7 serotypes are represented in the formulation. The novel vaccine
CC   of overcomes these problems, as it is easily purified, and available in
CC   large quantities. It is also expressed in the lymph nodes for a better
CC   immune response. The present sequence represents BoNTA heavy chain
CC   C-terminal subfragment AsubHc2, comprising residues 234 to 438 of the
CC   BoNTA Hc fragment plus an initila methionine, and was used in the present
CC   invention.
XX
SQ   Sequence   206 AA;

  Query Match           100.0%;  Score 727;  DB 21;  Length 206;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 119

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 179

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    180 QIERSSRTLGCSWEFIPVDD 199


RESULT 3
AAB36303
ID   AAB36303 standard; Protein; 382 AA.
XX
AC   AAB36303;
XX
DT   15-FEB-2001  (first entry)
XX
DE   BoNT/A prototoxin heavy chain C-terminal neural cell binding domain.
XX
KW   Human; procholecystokinin; CCK A receptor; CCK B receptor;
KW   pancreatitis; antiinflammatory.
XX
OS   Clostridium botulinum.
XX
PN   WO200061192-A2.
XX
PD   19-OCT-2000.
XX
PF   06-APR-2000; 2000WO-US09142.
XX
PR   08-APR-1999;   99US-0288326.
XX
PA   (ALLR ) ALLERGAN SALES INC.
XX
PI   Steward LE,  Sachs G,  Aoki KR;
XX
DR   WPI; 2000-679416/66.
XX
PT   New composition for treating acute pancreatitis, comprises a pancreatic
PT   cell surface marker binding element, a translocation element that
PT   transfers polypeptide across vesicular membrane, and a therapeutic
PT   element -
XX
PS   Example 1; Page 29; 50pp; English.
XX
```

```
CC   The present invention describes a composition (I) for treating acute
CC   pancreatitis. (I) comprises a first element containing a binding element
CC   that binds to a pancreatic cell surface marker, a second element
CC   containing a translocation element that facilitates polypeptide transfer
CC   across the vesicular membrane, and a third element containing a
CC   therapeutic element that inhibits enzyme secretion in pancreatic cell
CC   cytoplasm. Also described is a method for making a therapeutic
CC   polypeptide having a binding element selective for cholecystokinin (CCK)
CC   receptor by expressing within a host cell a recombinant chimeric
CC   polypeptide comprising an extein containing a therapeutic element and a
CC   translocational element, and an intein located to the carboxy terminal
CC   of extein having a cysteine, serine or threonine at its amino terminus,
CC   and contacting the extein with a synthetic peptide comprising a CCK
CC   amino acid sequence containing an amidated phenylalanine at a natural
CC   C-terminus, and a cysteine, serine or threonine at its N-terminus, and
CC   a nucleophilic reagent able to cause cleavage of the intein to form a
CC   peptide bond between the extein C-terminus and synthetic peptide
CC   N-terminus through the formation of an activated ester or thio ester
CC   intermediate. (I) has antiinflammatory activity and prevents accumulation
CC   of pancreatic digestive enzymes, and prevents exocytic fusion of vesicles
CC   containing secretory enzymes of pancreas. (I) is useful for treating
CC   acute pancreatitis. The present sequence represents the Clostridium
CC   botulinum BoNT/A prototoxin heavy chain C-terminal neural cell binding
CC   domain, which is given in the exemplification of the present invention.
XX
SQ   Sequence   382 AA;

  Query Match             100.0%;  Score 727;  DB 21;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     236 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 295

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     296 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 355

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db     356 QIERSSRTLGCSWEFIPVDD 375


RESULT 4
AAB04083
ID   AAB04083 standard; protein; 415 AA.
XX
AC   AAB04083;
XX
DT   11-APR-2001  (first entry)
XX
DE   Botulism toxin C fragment sequence (serotype A).
XX
KW   Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW   recombinant vector; antigen; immune response; vaccine; bacterium;
KW   infection.
XX
OS   Clostridium botulinum.
XX
PN   WO200067700-A2.
XX
PD   16-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US12890.
XX
PR   12-MAY-1999;   99US-0133865.
PR   12-MAY-1999;   99US-0133866.
PR   12-MAY-1999;   99US-0133867.
PR   12-MAY-1999;   99US-0133868.
PR   12-MAY-1999;   99US-0133869.
PR   12-MAY-1999;   99US-0133873.
PR   29-JUL-1999;   99US-0146192.
XX
PA   (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI   Smith LA,  Byrne MP,  Middlebrook JL,  Lapenotiere H;
XX
DR   WPI; 2001-016048/02.
XX
PT   New nucleic acids encoding the carboxy- or amino-terminal portions of
PT   the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT   vaccine against botulism
XX
PS   Example 7; Page 36; 73pp; English.
XX
CC   Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC   chain and then posttranslationally nicked, forming a dichain
CC   consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC   remain linked by a disulfide bond. Nucleic acids encoding the
CC   carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC   chain  of botulinum neurotoxin (BoNT) can be used in recombinant
CC   expression vectors and expressed in transformed cells to produce
CC   peptide antigens useful for eliciting an immune response to give
CC   protective immunity against botulinum neurotoxin, which causes
CC   botulism. The nucleic acids are expressible in a recombinant
CC   organisms such as Escherichia coli or Pichia pastoris. The use
CC   of recombinant nucleic acids are advantageous since it eliminates
CC   the need to culture large quantities of hazardous toxin-producing
CC   bacterium. Production yield from the genetically engineered product
CC   is also high and cost of production is lower. The nucleic acids can
CC   be derived from Clostridium botulinum serotypes A-G.
XX
SQ   Sequence   415 AA;

  Query Match             100.0%;  Score 727;  DB 22;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     269 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 328

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 388

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db     389 QIERSSRTLGCSWEFIPVDD 408


RESULT 5
AAY77142
ID   AAY77142 standard; Protein; 432 AA.
XX
AC   AAY77142;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Native botulinum neurotoxin serotype A (BoNTA) C-terminal fragment (Hc).
XX
KW   Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW   C-terminal fragment; Hc; Venezuelan equine encephalitis virus replicon;
KW   VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS   Clostridium botulinum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 432
FT                   /note= "Apparently encoded by GGATGGGGAG AAAGGCCACT G"
XX
PN   WO200002524-A2.
XX
PD   20-JAN-2000.
XX
```

PF 09-JUL-1999; 99WO-US15570.
XX
PR 10-JUL-1998; 98US-0092416.
PR 12-MAY-1999; 99US-0133870.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX
DR WPI; 2000-160827/14.
DR N-PSDB; AAZ87220.
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
PS Example 3; Page 52; 54pp; English.
XX
CC The invention relates to novel vaccines that induce a protective immune
CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC serotypes A-G. In preferred embodiments of the invention, the vector is a
CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC vector results in the production of large amounts of a protein encoded by
CC a sequence cloned into the replicon. The constructs are used to produce
CC vaccines against botulism. The proteins can also be used as diagnostic
CC tools for the diagnosis of botulism. The transformed host cells can be
CC used to analyse the effectiveness of drugs and agents which inhibit toxin
CC effects. The vaccine currently used against botulism is dangerous
CC and expensive to produce, and contains formalin, which is very painful
CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC the 7 serotypes are represented in the formulation. The novel vaccine
CC of overcomes these problems, as it is easily purified, and available in
CC large quantities. It is also expressed in the lymph nodes for a better
CC immune response. The present sequence represents the native BoNTA heavy
CC chain C-terminal fragment (Hc) used in an exemplification of the present
CC invention.
XX
SQ Sequence 432 AA;

Query Match 100.0%; Score 727; DB 21; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.7e-77;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 351

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 411

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    412 QIERSSRTLGCSWEFIPVDD 431
```

RESULT 6
AAB04089
ID AAB04089 standard; Protein; 434 AA.
XX
AC AAB04089;
XX
DT 11-APR-2001 (first entry)
XX
DE Botulism toxin heavy chain C-terminal sequence (serotype A).
XX
KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; immune response; vaccine; bacterium;
KW infection.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN WO200067700-A2.
XX
PD 16-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
XX
PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX
DR WPI; 2001-016048/02.
DR N-PSDB; AAA54483.
XX
PT New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulism
XX
PS Disclosure; Fig 2b; 73pp; English.
XX
CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.
XX
SQ Sequence 434 AA;

Query Match 100.0%; Score 727; DB 22; Length 434;
Best Local Similarity 100.0%; Pred. No. 6.7e-77;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 347

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    348 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 407

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    408 QIERSSRTLGCSWEFIPVDD 427
```

RESULT 7
AAB04090
ID AAB04090 standard; Protein; 435 AA.
XX
AC AAB04090;
XX
DT 11-APR-2001 (first entry)

```
XX
DE   Botulism toxin heavy chain C-terminal sequence (serotype A).
XX
KW   Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW   recombinant vector; antigen; immune response; vaccine; bacterium;
KW   infection.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO200067700-A2.
XX
PD   16-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US12890.
XX
PR   12-MAY-1999;   99US-0133865.
PR   12-MAY-1999;   99US-0133866.
PR   12-MAY-1999;   99US-0133867.
PR   12-MAY-1999;   99US-0133868.
PR   12-MAY-1999;   99US-0133869.
PR   12-MAY-1999;   99US-0133873.
PR   29-JUL-1999;   99US-0146192.
XX
PA   (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI   Smith LA,  Byrne MP,  Middlebrook JL,  Lapenotiere H;
XX
DR   WPI; 2001-016048/02.
DR   N-PSDB; AAA54484.
XX
PT   New nucleic acids encoding the carboxy- or amino-terminal portions of
PT   the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT   vaccine against botulism
XX
PS   Disclosure; Fig 3b; 73pp; English.
XX
CC   Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC   chain and then posttranslationally nicked, forming a dichain
CC   consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC   remain linked by a disulfide bond. Nucleic acids encoding the
CC   carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC   chain  of botulinum neurotoxin (BoNT) can be used in recombinant
CC   expression vectors and expressed in transformed cells to produce
CC   peptide antigens useful for eliciting an immune response to give
CC   protective immunity against botulinum neurotoxin, which causes
CC   botulism. The nucleic acids are expressible in a recombinant
CC   organisms such as Escherichia coli or Pichia pastoris. The use
CC   of recombinant nucleic acids are advantageous since it eliminates
CC   the need to culture large quantities of hazardous toxin-producing
CC   bacterium. Production yield from the genetically engineered product
CC   is also high and cost of production is lower. The nucleic acids can
CC   be derived from Clostridium botulinum serotypes A-G.
XX
SQ   Sequence   435 AA;

  Query Match             100.0%;  Score 727;  DB 22;  Length 435;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 348

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     349 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 408

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db     409 QIERSSRTLGCSWEFIPVDD 428
```

```
RESULT 8
AAB04088
ID   AAB04088 standard; Protein; 437 AA.
XX
AC   AAB04088;
XX
DT   11-APR-2001  (first entry)
XX
DE   Botulism toxin heavy chain C-terminal sequence (serotype A).
XX
KW   Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW   recombinant vector; antigen; immune response; vaccine; bacterium;
KW   infection.
XX
OS   Synthetic.
OS   Clostridium botulinum.
XX
PN   WO200067700-A2.
XX
PD   16-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US12890.
XX
PR   12-MAY-1999;   99US-0133865.
PR   12-MAY-1999;   99US-0133866.
PR   12-MAY-1999;   99US-0133867.
PR   12-MAY-1999;   99US-0133868.
PR   12-MAY-1999;   99US-0133869.
PR   12-MAY-1999;   99US-0133873.
PR   29-JUL-1999;   99US-0146192.
XX
PA   (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
PI   Smith LA,  Byrne MP,  Middlebrook JL,  Lapenotiere H;
XX
DR   WPI; 2001-016048/02.
DR   N-PSDB; AAA54482.
XX
PT   New nucleic acids encoding the carboxy- or amino-terminal portions of
PT   the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT   vaccine against botulism
XX
PS   Claim 3; Fig 1b; 73pp; English.
XX
CC   Botulism neurotoxins are translated as a single 150 kDa polypeptide
CC   chain and then posttranslationally nicked, forming a dichain
CC   consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC   remain linked by a disulfide bond. Nucleic acids encoding the
CC   carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC   chain  of botulinum neurotoxin (BoNT) can be used in recombinant
CC   expression vectors and expressed in transformed cells to produce
CC   peptide antigens useful for eliciting an immune response to give
CC   protective immunity against botulinum neurotoxin, which causes
CC   botulism. The nucleic acids are expressible in a recombinant
CC   organisms such as Escherichia coli or Pichia pastoris. The use
CC   of recombinant nucleic acids are advantageous since it eliminates
CC   the need to culture large quantities of hazardous toxin-producing
CC   bacterium. Production yield from the genetically engineered product
CC   is also high and cost of production is lower. The nucleic acids can
CC   be derived from Clostridium botulinum serotypes A-G.
XX
SQ   Sequence   437 AA;

  Query Match             100.0%;  Score 727;  DB 22;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 350

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    351 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 410

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    411 QIERSSRTLGCSWEFIPVDD 430


RESULT 9
AAR95008
ID   AAR95008 standard; Protein; 438 AA.
XX
AC   AAR95008;
XX
DT   07-JUL-1996  (first entry)
XX
DE   Type A neurotoxin C fragment.
XX
KW   Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;
KW   Clostridium botulinum.
XX
OS   Synthetic.
XX
PN   WO9612802-A1.
XX
PD   02-MAY-1996.
XX
PF   23-OCT-1995;   95WO-US13737.
XX
PR   07-JUN-1995;   95US-0480604.
PR   24-OCT-1994;   94US-0329154.
PR   16-MAR-1995;   95US-0405496.
PR   14-APR-1995;   95US-0422711.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Firca JR,  Kink JA,  Padhye NV,  Stafford DC,  Thalley BS;
PI   Williams JA;
XX
DR   WPI; 1996-230603/23.
DR   N-PSDB; AAT29245.
XX
PT   Fusion proteins comprising non-toxin protein and part of toxin  -
PT   useful to form anti-toxins against Clostridium botulinum type A, and
PT   C. difficile type toxins, and to treat C. difficile intoxication,
PT   partic. diarrhoea
XX
PS   Claim 5; Page 336-38; 434pp; English.
XX
CC   The heavy chain C fragment (AAR95008) of Clostridium botulinum type
CC   A neurotoxin (see also AAR95010) was produced by expression of a
CC   synthetic gene (AAT29245) in Escherichia coli.  The C fragment
CC   comprises the C-terminal portion of the type A neurotoxin heavy
CC   chain.  It is pref. produced as a fusion protein with a
CC   polyhistidine affinity tag or maltose binding protein, and is
CC   used to produce avian antitoxin against C. botulinum type A
CC   or as an immunogen in vaccine compsns. (see also AAR95009).
XX
SQ   Sequence   438 AA;

  Query Match            100.0%;  Score 727;  DB 17;  Length 438;
  Best Local Similarity  100.0%;  Pred. No. 6.8e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 351

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 411

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    412 QIERSSRTLGCSWEFIPVDD 431


RESULT 10
AAW68389
ID   AAW68389 standard; Protein; 438 AA.
XX
AC   AAW68389;
XX
DT   07-DEC-1998  (first entry)
XX
DE   Clostridium botulinum toxin A fragment C.
XX
KW   Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
KW   botulism.
XX
OS   Clostridium botulinum serotype A.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..2
FT                   /note= "vector-derived amino acid residues"
XX
PN   WO9808540-A1.
XX
PD   05-MAR-1998.
XX
PF   28-AUG-1997;   97WO-US15394.
XX
PR   28-AUG-1996;   96US-0704159.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Thalley BS,  Williams JA;
XX
DR   WPI; 1998-230234/20.
DR   N-PSDB; AAV30571.
XX
PT   Host cell containing recombinant expression vector encoding
PT   Clostridium botulinum type B or E toxin  - useful to treat humans
PT   and other animals at risk of intoxication with clostridial toxin
XX
PS   Example 22; Page 262-263; 428pp; English.
XX
CC   This is the amino acid sequence of Clostridium botulinum serotype A
CC   toxin C-fragment expressed by a DNA sequence (see AAV30571) in plasmid
CC   pAlterBot.  Recombinant C-fragment proteins have been produced in
CC   Escherichia coli as fusion proteins with either maltose binding
CC   protein or Clostridium difficile type A toxin (see AAW68387).  The
CC   invention relates to recombinant proteins derived from C. botulinum
CC   toxins.  Methods are provided which allow for the isolation of
CC   soluble recombinant toxin proteins free of significant endotoxin
CC   contamination.  Preferred hosts for production of the recombinant
CC   proteins are E. coli, insect cells and yeast cells.  The recombinant
CC   toxin proteins are used as immunogens for the production of vaccines
CC   and antitoxins that are useful in the treatment of humans and
CC   animals at risk of intoxication with clostridial toxin.
XX
SQ   Sequence   438 AA;

  Query Match            100.0%;  Score 727;  DB 19;  Length 438;
  Best Local Similarity  100.0%;  Pred. No. 6.8e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 351

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 411

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
```

```
Db    412 QIERSSRTLGCSWEFIPVDD 431


RESULT 11
AAY77134
ID   AAY77134 standard; Protein; 438 AA.
XX
AC   AAY77134;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Synthetic botulinum neurotoxin serotype A (BoNTA) C-terminal fragment.
XX
KW   Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW   C-terminal fragment; Venezuelan equine encephalitis virus replicon;
KW   VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS   Clostridium botulinum.
OS   Synthetic.
XX
PN   WO200002524-A2.
XX
PD   20-JAN-2000.
XX
PF   09-JUL-1999;   99WO-US15570.
XX
PR   10-JUL-1998;   98US-0092416.
PR   12-MAY-1999;   99US-0133870.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Lee JS,  Pushko P,  Smith JF,  Parker M,  Dertzbaugh MT,  Smith L;
XX
DR   WPI; 2000-160827/14.
DR   N-PSDB; AAZ87212.
XX
PT   Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT   toxin serotypes A-G, is used for inducing an immune response against
PT   botulinum -
XX
PS   Claim 22; Page 54; 54pp; English.
XX
CC   The invention relates to novel vaccines that induce a protective immune
CC   response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC   and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC   DNA construct comprising a vector, and at least one nucleic acid
CC   fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC   serotypes A-G. In preferred embodiments of the invention, the vector is
CC   a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of
CC   this vector results in the production of large amounts of a protein
CC   encoded by a sequence cloned into the replicon. The constructs are used
CC   to produce vaccines against botulism. The proteins can also be used as
CC   diagnostic tools for the diagnosis of botulism. The transformed host
CC   cells can be used to analyse the effectiveness of drugs and agents which
CC   inhibit toxin effects. The vaccine currently used against botulism is
CC   dangerous and expensive to produce, and contains formalin, which is very
CC   painful for the recipient. Also, the vaccine is incomplete, in that only
CC   5 of the 7 serotypes are represented in the formulation. The novel
CC   vaccine of overcomes these problems, as it is easily purified, and
CC   available in large quantities. It is also expressed in the lymph nodes
CC   for a better immune response. Sequences AAY77134-Y77139 represent
CC   synthetic BoNT Hc fragments used in the present invention. The DNA
CC   encoding these sequences had been optimised for codon usage for
CC   expression in yeast.
XX
SQ   Sequence   438 AA;

  Query Match            100.0%;  Score 727;  DB 21;  Length 438;
  Best Local Similarity  100.0%;  Pred. No. 6.8e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    292 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 351

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 411

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    412 QIERSSRTLGCSWEFIPVDD 431


RESULT 12
AAW68391
ID   AAW68391 standard; Protein; 445 AA.
XX
AC   AAW68391;
XX
DT   07-DEC-1998  (first entry)
XX
DE   Clostridium botulinum toxin A fragment C (His-tagged).
XX
KW   Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
KW   botulism.
XX
OS   Clostridium botulinum serotype A.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..7
FT                   /note= "N-terminal histidine tag"
XX
PN   WO9808540-A1.
XX
PD   05-MAR-1998.
XX
PF   28-AUG-1997;   97WO-US15394.
XX
PR   28-AUG-1996;   96US-0704159.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Thalley BS,  Williams JA;
XX
DR   WPI; 1998-230234/20.
DR   N-PSDB; AAV30576.
XX
PT   Host cell containing recombinant expression vector encoding
PT   Clostridium botulinum type B or E toxin  - useful to treat humans
PT   and other animals at risk of intoxication with clostridial toxin
XX
PS   Example 29; Page 279-281; 428pp; English.
XX
CC   This is the amino acid sequence of a histidine-tagged fragment C
CC   polypeptide of Clostridium botulinum serotype A toxin encoded by a
CC   DNA sequence (see AAV30576) in plasmid p6HisBot(syn).  This vector
CC   was used to express native soluble C fragment in Escherichia coli
CC   host cells, with the recombinant C fragment being purified on a
CC   poly-histidine binding affinity resin.  The invention relates
CC   to recombinant proteins derived from C. botulinum toxins.  Methods
CC   are provided which allow the isolation of soluble recombinant
CC   proteins that are free of significant endotoxin contamination.
CC   Preferred hosts for production of recombinant proteins are E. coli,
CC   insect cells and yeast cells.  The recombinant toxins are used as
CC   immunogens for the production of vaccines and antitoxins that are
CC   useful in the treatment of humans and animals at risk of
CC   intoxication with clostridial toxin.
XX
SQ   Sequence   445 AA;

  Query Match            100.0%;  Score 727;  DB 19;  Length 445;
  Best Local Similarity  100.0%;  Pred. No. 7e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 358

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 418

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    419 QIERSSRTLGCSWEFIPVDD 438


RESULT 13
AAR95009
ID   AAR95009 standard; Protein; 462 AA.
XX
AC   AAR95009;
XX
DT   07-JUL-1996  (first entry)
XX
DE   Type A neurotoxin C fragment-polyhistidine tag fusion pHisBot.
XX
KW   Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;
KW   Clostridium botulinum; polyhistidine; vector; pETHisa; pHisBot.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..21
FT                   /label= Polyhistidine_tag
FT   Protein         22..462
FT                   /label= C-fragment
XX
PN   WO9612802-A1.
XX
PD   02-MAY-1996.
XX
PF   23-OCT-1995;   95WO-US13737.
XX
PR   07-JUN-1995;   95US-0480604.
PR   24-OCT-1994;   94US-0329154.
PR   16-MAR-1995;   95US-0405496.
PR   14-APR-1995;   95US-0422711.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Firca JR,  Kink JA,  Padhye NV,  Stafford DC,  Thalley BS;
PI   Williams JA;
XX
DR   WPI; 1996-230603/23.
DR   N-PSDB; AAT29246.
XX
PT   Fusion proteins comprising non-toxin protein and part of toxin -
PT   useful to form anti-toxins against Clostridium botulinum type A, and
PT   C. difficile type toxins, and to treat C. difficile intoxication,
PT   partic. diarrhoea
XX
PS   Claim 7; Page 340-342; 434pp; English.
XX
CC   pHisBot fusion protein (AAR95009), the product of a nucleotide
CC   sequence (AAT29246) in vector pETHisa, comprises a polyhistidine
CC   affinity tag and fragment C (see also AAR95008) of the Clostridium
CC   botulinum type A neurotoxin.  The pHisBot protein was expressed
CC   in Escherichia coli as a soluble protein and was purified by
CC   metal chelate affinity chromatography to obtain a product free
CC   of endotoxin contamination that may be useful as an immunogen
CC   in vaccine compsns.
XX
SQ   Sequence   462 AA;

  Query Match           100.0%;  Score 727;  DB 17;  Length 462;
  Best Local Similarity  100.0%;  Pred. No. 7.4e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 375

Qy     61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 435

Qy    121 QIERSSRTLGCSWEFIPVDD 140
          ||||||||||||||||||||
Db    436 QIERSSRTLGCSWEFIPVDD 455


RESULT 14
AAW68390
ID   AAW68390 standard; Protein; 462 AA.
XX
AC   AAW68390;
XX
DT   07-DEC-1998  (first entry)
XX
DE   Clostridium botulinum toxin A fragment C (His-tagged).
XX
KW   Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;
KW   botulism.
XX
OS   Clostridium botulinum serotype A.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..21
FT                   /note= "N-terminal histidine tag"
XX
PN   WO9808540-A1.
XX
PD   05-MAR-1998.
XX
PF   28-AUG-1997;   97WO-US15394.
XX
PR   28-AUG-1996;   96US-0704159.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Thalley BS,  Williams JA;
XX
DR   WPI; 1998-230234/20.
DR   N-PSDB; AAV30572 and AAV30575.
XX
PT   Host cell containing recombinant expression vector encoding
PT   Clostridium botulinum type B or E toxin  - useful to treat humans
PT   and other animals at risk of intoxication with clostridial toxin
XX
PS   Example 24; Page 265-267; 428pp; English.
XX
CC   This is the amino acid sequence of a histidine-tagged fragment C
CC   polypeptide of Clostridium botulinum serotype A toxin encoded by a
CC   DNA sequence (see AAV30572) in plasmid pHisBot, and by a DNA sequence
CC   (see AAV30575) in pHisBotA.  These vectors were used to express
CC   native (i.e. non-fusion) soluble C fragment in Escherichia coli
CC   host cells, with the recombinant C fragment being purified on a
CC   poly-histidine binding affinity resin.  The invention relates
CC   to recombinant proteins derived from C. botulinum toxins.  Methods
CC   are provided which allow the isolation of soluble recombinant
CC   proteins that are free of significant endotoxin contamination.
CC   Preferred hosts for production of recombinant proteins are E. coli,
CC   insect cells and yeast cells.  The recombinant toxins are used as
CC   immunogens for the production of vaccines and antitoxins that are
CC   useful in the treatment of humans and animals at risk of
CC   intoxication with clostridial toxin.
XX
SQ   Sequence   462 AA;
```

```
  Query Match             100.0%;  Score 727;  DB 19;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-77;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 375

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 435

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db     436 QIERSSRTLGCSWEFIPVDD 455


RESULT 15
AAY77140
ID   AAY77140 standard; Protein; 837 AA.
XX
AC   AAY77140;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Native botulinum neurotoxin serotype A (BoNTA).
XX
KW   Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW   Venezuelan equine encephalitis virus replicon;
KW   VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS   Clostridium botulinum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 837
FT                   /note= "Apparently encoded by GGATGGGGAG AAAGGCCACT G"
XX
PN   WO200002524-A2.
XX
PD   20-JAN-2000.
XX
PF   09-JUL-1999;   99WO-US15570.
XX
PR   10-JUL-1998;   98US-0092416.
PR   12-MAY-1999;   99US-0133870.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Lee JS,  Pushko P,  Smith JF,  Parker M,  Dertzbaugh MT,  Smith L;
XX
DR   WPI; 2000-160827/14.
DR   N-PSDB; AAZ87218.
XX
PT   Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT   toxin serotypes A-G, is used for inducing an immune response against
PT   botulinum -
XX
PS   Example 3; Page 49; 54pp; English.
XX
CC   The invention relates to novel vaccines that induce a protective immune
CC   response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC   and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC   DNA construct comprising a vector, and at least one nucleic acid
CC   fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC   serotypes A-G. In preferred embodiments of the invention, the vector is a
CC   Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC   vector results in the production of large amounts of a protein encoded by
CC   a sequence cloned into the replicon. The constructs are used to produce
CC   vaccines against botulism. The proteins can also be used as diagnostic
CC   tools for the diagnosis of botulism. The transformed host cells can be
CC   used to analyse the effectiveness of drugs and agents which inhibit toxin
CC   effects. The vaccine currently used against botulism is dangerous
CC   and expensive to produce, and contains formalin, which is very painful
CC   for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC   the 7 serotypes are represented in the formulation. The novel vaccine
CC   of overcomes these problems, as it is easily purified, and available in
CC   large quantities. It is also expressed in the lymph nodes for a better
CC   immune response. The present sequence represents the native BoNTA heavy
CC   chain used in an exemplification of the present invention.
XX
SQ   Sequence   837 AA;

  Query Match             100.0%;  Score 727;  DB 21;  Length 837;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-76;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     697 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 756

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     757 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 816

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db     817 QIERSSRTLGCSWEFIPVDD 836


Search completed: March 13, 2003, 11:39:09
Job time : 26.2248 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:36:37 ; Search time 6.45533 Seconds
(without alignments)
899.518 Million cell updates/sec

Title: US-09-917-791-22
Perfect score: 727
Sequence: 1 LNSSLYRGTKFIIKKYASGN.....QIERSSRTLGCSWEFIPVDD 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	727	100.0	1295	1	BXA1_CLOBO	P10845 clostridium
2	671	92.3	1295	1	BXA2_CLOBO	Q45894 clostridium
3	258.5	35.6	1274	1	BXF_CLOBO	P30996 clostridium
4	199	27.4	1250	1	BXE_CLOBO	Q00496 clostridium
5	190	26.1	1250	1	BXE_CLOBU	P30995 clostridium
6	165.5	22.8	1290	1	BXB_CLOBO	P10844 clostridium
7	163	22.4	1314	1	TETX_CLOTE	P04958 clostridium
8	158.5	21.8	1296	1	BXG_CLOBO	Q60393 clostridium
9	100	13.8	1290	1	BXC1_CLOBO	P18640 clostridium
10	88.5	12.2	264	1	MURI_FUSNN	Q8ree6 fusobacteri
11	79.5	10.9	1786	1	YCF1_ARATH	P56785 arabidopsis
12	79	10.9	765	1	Y008_HUMAN	Q15398 homo sapien
13	75	10.3	720	1	WZC_ECOLI	P76387 escherichia
14	74	10.2	1276	1	BXD_CLOBO	P19321 clostridium
15	73.5	10.1	682	1	NISP_LACLA	Q07596 lactococcus
16	73	10.0	556	1	MUTL_SYNY3	P73349 synechocyst
17	73	10.0	573	1	TLPC_BACSU	P39209 bacillus su
18	73	10.0	720	1	WZC_ECO57	Q8x719 escherichia
19	72.5	10.0	875	1	ZIP1_YEAST	P31111 saccharomyc
20	72	9.9	1177	1	Y307_MYCGE	P47549 mycoplasma
21	71.5	9.8	399	1	LOLC_BUCAI	P57382 buchnera ap
22	71.5	9.8	721	1	PRTP_HSV7J	P52385 human herpe
23	71.5	9.8	735	1	CIGB_DICDI	Q94481 dictyosteli
24	71.5	9.8	1076	1	RPOB_ASTLO	P27059 astasia lon
25	71.5	9.8	1744	1	TANA_XENLA	Q01550 xenopus lae
26	71.5	9.8	2339	1	RPC1_PLAFA	P27625 plasmodium
27	71	9.8	232	1	MTN_BUCAI	P57306 buchnera ap
28	71	9.8	266	1	TERM_BPPH2	P03681 bacteriopha
29	71	9.8	353	1	RPOA_MYCSP	P38018 mycoplasma
30	71	9.8	445	1	II2A_MOUSE	P15091 mus musculu
31	71	9.8	445	1	II2B_MOUSE	Q9r002 mus musculu
32	71	9.8	542	1	V155_FOWPV	Q9j5a7 fowlpox vir
33	71	9.8	3079	1	IRA2_YEAST	P19158 saccharomyc
34	70.5	9.7	607	1	YMH1_YEAST	Q04632 saccharomyc
35	70.5	9.7	831	1	IF2_RICPR	Q9zcz8 rickettsia
36	70.5	9.7	926	1	UBP4_YEAST	P32571 saccharomyc
37	70.5	9.7	964	1	PMPE_CHLTR	O84877 chlamydia t
38	70.5	9.7	1089	1	NMD2_YEAST	P38798 saccharomyc
39	70.5	9.7	1325	1	YDEK_ECOLI	P32051 escherichia
40	70	9.6	145	1	RL9_UREPA	Q9ppt9 ureaplasma
41	70	9.6	167	1	YREP_BUCAP	O85061 buchnera ap
42	70	9.6	359	1	TMOD_HUMAN	P28289 homo sapien
43	70	9.6	1014	1	NANH_CLOSE	P29767 clostridium
44	70	9.6	4725	1	DYHC_DICDI	P34036 dictyosteli
45	69.5	9.6	199	1	GRPE_LEGPN	O32481 legionella

ALIGNMENTS

RESULT 1
BXA1_CLOBO
ID BXA1_CLOBO STANDARD; PRT; 1295 AA.
AC P10845; P18639; P01561;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 2916;
RX MEDLINE=90235864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
RA Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT gene.";
RL Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=62A;
RX MEDLINE=90264400; PubMed=2160960;
RA Binz B., Kuarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=62A;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=Hall;
RX MEDLINE=89350959; PubMed=2669749;
RA Betley M.J., Somers E., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
RT the N-terminal encoding region.";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=Type A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;

```
RT   "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT   components of Clostridium botulinum type A progenitor toxins.";
RL   FEBS Lett. 376:41-44(1995).
RN   [6]
RP   SEQUENCE OF 1-16.
RX   MEDLINE=84178501; PubMed=6370252;
RA   Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequence of the heavy and light chains of
RT   botulinum neurotoxin type A.";
RL   Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN   [7]
RP   SEQUENCE OF 1-46.
RA   Dasgupta B.R., Foley J., Niece R.;
RT   "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL   Biochemistry 26:4162-4162(1987).
RN   [8]
RP   SEQUENCE OF 1-5 AND 444-456.
RX   MEDLINE=91120847; PubMed=2126206;
RA   Dasgupta B.R., Dekleva M.L.;
RT   "Botulinum neurotoxin type A: sequence of amino acids at the
RT   N-terminus and around the nicking site.";
RL   Biochimie 72:661-664(1990).
RN   [9]
RP   SEQUENCE OF 448-464 AND 872-895.
RX   MEDLINE=89024662; PubMed=3178218;
RA   Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT   "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RT   halves and their partial sequences.";
RL   Arch. Biochem. Biophys. 266:142-151(1988).
RN   [10]
RP   SEQUENCE OF 448-482.
RX   MEDLINE=85285016; PubMed=3896784;
RA   Shone C.C., Hambleton P., Melling J.;
RT   "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RT   and purification of two tryptic fragments. Proteolytic action near
RT   the COOH-terminus of the heavy subunit destroys toxin-binding
RT   activity.";
RL   Eur. J. Biochem. 151:75-82(1985).
RN   [11]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [12]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
RN   [13]
RP   MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX   MEDLINE=21556941; PubMed=11700044;
RA   Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
RT   "Site-directed mutagenesis identifies active-site residues of the
RT   light chain of botulinum neurotoxin type a.";
RL   Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX   MEDLINE=98455071; PubMed=9783750;
RA   Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT   "Crystal structure of botulinum neurotoxin type A and implications
RT   for toxicity.";
RL   Nat. Struct. Biol. 5:898-902(1998).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
CC       the treatment of strabismus and blepharospasm associated with
CC       dystonia and cervical dystonia. Also used for the treatment of
CC       hemifacial spasm and a number of other neurological disorders
CC       characterized by abnormal muscle contraction.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   -!- DATABASE: NAME=BOTOX product information Web site;
CC       WWW="http://www.botox.com/index.jsp?hp&productinfo".
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 19 of February 2002;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt019.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52066; CAA36289.1; -.
DR   EMBL; M30196; AAA23262.1; -.
DR   EMBL; X92973; CAA63551.1; -.
DR   EMBL; D67030; BAA11051.1; -.
DR   EMBL; M27892; AAA23269.1; -.
DR   PIR; A35294; BTCLAB.
DR   PIR; S09492; S09492.
DR   PDB; 3BTA; 01-OCT-99.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   Pharmaceutical; 3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC).
FT   ACT_SITE    223    223
FT   METAL       226    226       ZINC (CATALYTIC).
FT   METAL       261    261       ZINC (CATALYTIC).
FT   DISULFID    429    453       INTERCHAIN.
FT   DISULFID   1234   1279
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
FT   VARIANT      26     26       V -> A.
FT   MUTAGEN     261    261       E->A: DRASTIC DECREASE IN ENZYMATIC
FT                                ACTIVITY.
FT   MUTAGEN     265    265       F->A: DECREASE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN     365    365       Y->A: DECREASE IN ENZYMATIC ACTIVITY.
FT   CONFLICT      1      1       P -> Q (IN REF. 1).
FT   CONFLICT    479    479       E -> P (IN REF. 9).
FT   CONFLICT    875    875       T -> L (IN REF. 8).
FT   CONFLICT    891    891       S -> K (IN REF. 8).
SQ   SEQUENCE   1295 AA;  149322 MW;  858342F754862579 CRC64;

  Query Match           100.0%;  Score 727;  DB 1;  Length 1295;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.8e-59;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1149 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 1208

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1209 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 1268

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db    1269 QIERSSRTLGCSWEFIPVDD 1288


RESULT 2
BXA2_CLOBO
ID   BXA2_CLOBO     STANDARD;      PRT;  1295 AA.
AC   Q45894; P77780;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=94143603; PubMed=8310180;
RA   Willems A., East A.K., Lawson P.A., Collins M.D.;
RT   "Sequence of the gene coding for the neurotoxin of Clostridium
RT   botulinum type A associated with infant botulism: comparison with
RT   other clostridial neurotoxins.";
RL   Res. Microbiol. 144:547-556(1993).
RN   [2]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure (By similarity).
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73423; CAA51824.1; -.
DR   EMBL; X87974; CAA61234.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    223    223       BY SIMILARITY.
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    453       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   1234   1279       BY SIMILARITY.
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
SQ   SEQUENCE   1295 AA;  149279 MW;  5DA04A13D98D6372 CRC64;

  Query Match             92.3%;  Score 671;  DB 1;  Length 1295;
  Best Local Similarity   90.7%;  Pred. No. 4.1e-54;
  Matches  127;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           |||:|| |||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db    1149 LNSTLYEGTKFIIKKYASGNEDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 1208

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           |||||||||||||||||:|||| |||||||||||||||||| ::|||||||||||||||
Db    1209 EIPDVGNLSQVVVMKSKDDQGIRNKCKMNLQDNNGNDIGFIGFHLYDNIAKLVASNWYNR 1268

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           |: ::||| ||||||||||
Db    1269 QVGKASRTFGCSWEFIPVDD 1288


RESULT 3
BXF_CLOBO
ID   BXF_CLOBO      STANDARD;      PRT;  1274 AA.
AC   P30996;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE   (Bontoxilysin F).
GN   BOTF.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23387;
RX   MEDLINE=93012902; PubMed=1398040;
RA   East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA   Roberts T.A., Thompson D.E.;
RT   "Sequence of the gene encoding type F neurotoxin of Clostridium
RT   botulinum.";
RL   FEMS Microbiol. Lett. 75:225-230(1992).
RN   [2]
RP   SEQUENCE OF 1-64 FROM N.A.
RC   STRAIN=Hobbs FT10;
RX   MEDLINE=94297488; PubMed=7764998;
RA   East A.K., Collins M.D.;
```

```
RT   "Conserved structure of genes encoding components of botulinum
RT   neurotoxin complex M and the sequence of the gene coding for the
RT   nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL   Curr. Microbiol. 29:69-77(1994).
RN   [3]
RP   SEQUENCE OF 634-1002 FROM N.A.
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
CC       BOND OF SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92906; AAA23263.1; -.
DR   EMBL; S73676; AAC60475.1; -.
DR   EMBL; X70820; CAA50151.1; -.
DR   EMBL; X70816; CAA50147.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    436       BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT   CHAIN       437   1274       BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT   METAL       227    227       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    228    228       BY SIMILARITY.
FT   METAL       231    231       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    445       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1274 AA;  146709 MW;  5B99756A7438B921 CRC64;

  Query Match             35.6%;  Score 258.5;  DB 1;  Length 1274;
  Best Local Similarity   43.2%;  Pred. No. 4.4e-16;
  Matches   60;  Conservative   20;  Mismatches   44;  Indels   15;  Gaps    5;

Qy       1 LNSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKIL 57
           ||  || | : ||:|       | || || ||  ||||| :  ||||  : :::  |||:

Db    1140 LNYKLYEGVEVIIRKNGPIDISNTDNFVRKNDLAYINVVVDRGVEYRLYAD-TKSEKEKII 1198

Qy      58 SALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
                : |  :| |::|| |     | | | || |:|||::|| :||| |    ||||:|
Db    1199 RTSNLND--SLGQIIVMDS-----IGNNCTMNFQNNNGSNIGLLGFHSNN----LVASSW 1247

Qy     118 YNRQIERSSRTLGCSWEFI 136
           |   | |:: : || |  |
Db    1248 YYNNIRRNTSSNGCFWSSI 1266


RESULT 4
BXE_CLOBO
ID   BXE_CLOBO      STANDARD;      PRT;  1250 AA.
AC   Q00496;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beluga;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92174922; PubMed=1541280;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RT   "The complete amino acid sequence of the Clostridium botulinum type-E
RT   neurotoxin, derived by nucleotide-sequence analysis of the encoding
RT   gene.";
RL   Eur. J. Biochem. 204:657-667(1992).
RN   [3]
RP   SEQUENCE OF 1-251 FROM N.A.
RX   MEDLINE=90264400; PubMed=2160960;
RA   Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
RT   "The complete sequence of botulinum neurotoxin type A and comparison
RT   with other clostridial neurotoxins.";
RL   J. Biol. Chem. 265:9153-9158(1990).
RN   [4]
RP   SEQUENCE OF 1-13.
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [5]
RP   SEQUENCE OF 419-426.
RX   MEDLINE=90344918; PubMed=2116911;
RA   Gimenez J.A., Dasgupta B.R.;
RT   "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT   reveals the site trypsin nicks and homology with tetanus
RT   neurotoxin.";
RL   Biochimie 72:213-217(1990).
RN   [6]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
```

```
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC       181 BOND IN SNAP-25.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62089; CAA43999.1; -.
DR   EMBL; X62683; CAA44558.1; -.
DR   PIR; A60027; A60027.
DR   PIR; B35294; B35294.
DR   PIR; JH0257; JH0257.
DR   PIR; S08575; S08575.
DR   PIR; S18111; S18111.
DR   PIR; S21178; S21178.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    176    176       R -> G (IN REF. 2).
FT   CONFLICT    197    197       C -> S (IN REF. 2 AND 3).
FT   CONFLICT    339    339       R -> A (IN REF. 2).
FT   CONFLICT    772    772       I -> L (IN REF. 2).
FT   CONFLICT    962    963       FE -> LQ (IN REF. 2).
FT   CONFLICT    966    966       R -> A (IN REF. 2).
FT   CONFLICT   1194   1194       N -> NN (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143712 MW;  D9FCE26DDA041EB4 CRC64;

  Query Match             27.4%;  Score 199;  DB 1;  Length 1250;
  Best Local Similarity   37.4%;  Pred. No. 1.3e-10;
  Matches   52;  Conservative   19;  Mismatches   52;  Indels   16;  Gaps    6;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVV-KNKEYRLATNASQAGVEKILS 58
           | : || | |  |::   |   ||:|| ||:|||| |  |   : |  : :    ||
Db    1116 LANRLYSGIKVKIQRVNNSSTNDNLVRKNDQVYINFVASKTHLFPLYADTATTNKEK--- 1172

Qy      59 ALEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
           ::|   ||  :|||| |         | || ::||||:|| :||       :||| |
Db    1173 TIKISSSGNRFNQVVVMNS------VGNCTMNFKNNNGNNIGLLGF----KADTVVASTW 1222

Qy     118 YNRQIERSSRTLGCSWEFI 136
           |   :   : : || | ||
Db    1223 YYTHMRDHTNSNGCFWNFI 1241


RESULT 5
BXE_CLOBU
ID   BXE_CLOBU      STANDARD;      PRT;  1250 AA.
AC   P30995;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 43181, and ATCC 43755;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE OF 1-251 FROM N.A.
RC   STRAIN=BL6340;
RX   MEDLINE=91237316; PubMed=2033376;
RA   Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA   Yokosawa N., Yashiki T., Oguma K.;
RT   "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT   type E toxin gene from Clostridium butyricum strain BL6340.";
RL   J. Gen. Microbiol. 137:519-525(1991).
RN   [3]
RP   SEQUENCE OF 1-48.
RC   STRAIN=5262;
RA   Gimenez J., Foley J., Dasgupta B.R.;
RT   "Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT   partial sequence and comparison.";
RL   FASEB J. 2:A1750-A1750(1988).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X62088; CAA43998.1; -.
DR   EMBL; X53180; CAA37321.1; -.
DR   PIR; JH0256; JH0256.
DR   PIR; S16145; S16145.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    229    229       K -> M (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143265 MW;  8171B5B2C2312857 CRC64;

  Query Match             26.1%;  Score 190;  DB 1;  Length 1250;
  Best Local Similarity   37.7%;  Pred. No. 8.9e-10;
  Matches   52;  Conservative   19;  Mismatches   53;  Indels   14;  Gaps    6;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSA 59
           | : || | |  |::   |   ||:|| ||:|||| |  :| : |   |  |   |
Db    1116 LANRLYSGIKVKIQRVNNSSTNDNLVRKNDQVYINFVA-SKTHLLPLYADTATTNK-EKT 1173

Qy      60 LEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWY 118
           ::|   ||  :|||| |          | || ::|||:|| :||        :||| ||
Db    1174 IKISSSGNRFNQVVVMNS------VGNCTMNFKNNNGNNIGLLGF----KADTVVASTWY 1223

Qy     119 NRQIERSSRTLGCSWEFI 136
             :  :: : |  | ||
Db    1224 YTHMRDNTNSNGFFWNFI 1241


RESULT 6
BXB_CLOBO
ID   BXB_CLOBO      STANDARD;      PRT;  1290 AA.
AC   P10844; P10843;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE   (Bontoxilysin B).
GN   BOTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92384550; PubMed=1514783;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA   Minton N.P.;
RT   "Molecular cloning of the Clostridium botulinum structural gene
RT   encoding the type B neurotoxin and determination of its entire
RT   nucleotide sequence.";
RL   Appl. Environ. Microbiol. 58:2345-2354(1992).
RN   [2]
RP   SEQUENCE OF 35-245 FROM N.A.
RC   STRAIN=NCTC 7273;
RA   Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 633-993 FROM N.A.
RC   STRAIN=NCTC 7273;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-44 AND 441-466.
RC   STRAIN=657;
RX   MEDLINE=89000987; PubMed=3139097;
RA   Dasgupta B.R., Datta A.;
RT   "Botulinum neurotoxin type B (strain 657): partial sequence and
RT   similarity with tetanus toxin.";
RL   Biochimie 70:811-817(1988).
RN   [5]
RP   SEQUENCE OF 1-16 AND 441-458.
RC   STRAIN=OKRA;
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93054694; PubMed=1429690;
RA   Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT   "Botulinum neurotoxins are zinc proteins.";
RL   J. Biol. Chem. 267:23479-23483(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC       SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81186; AAA23211.1; -.
DR   EMBL; Z11934; CAA77991.1; -.
DR   EMBL; X70817; CAA50148.1; -.
DR   PIR; S07128; S07128.
DR   PIR; S07155; S07155.
DR   PIR; S08562; S08562.
DR   PIR; S08573; S08573.
DR   PIR; S08574; S08574.
DR   PIR; A48940; A48940.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
```

```
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    440       BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT   CHAIN       441   1290       BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    445       INTERCHAIN (PROBABLE).
FT   CONFLICT     29     29       T -> M (IN REF. 4).
FT   CONFLICT    217    217       R -> G (IN REF. 2).
FT   CONFLICT    224    224       A -> S (IN REF. 2).
FT   CONFLICT    463    463       S -> R (IN REF. 4).
SQ   SEQUENCE   1290 AA;  150670 MW;  D21746E2C024DF43 CRC64;

  Query Match             22.8%;  Score 165.5;  DB 1;  Length 1290;
  Best Local Similarity   28.7%;  Pred. No. 1.7e-07;
  Matches   43;  Conservative   31;  Mismatches   61;  Indels   15;  Gaps    5;

Qy       5 LYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEI 62
           || | ||||:: ::     |:|||  | :|::    |:|:|: |       |: |    |
Db    1138 LYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNLNQEWRVYTYKYFKKEEEKLFLAPI 1197

Qy      63 PDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIA--------KL 112
             |       : :|  ::|  |  |::     : : ::|| || |:|
Db    1198 SDSDEFYNTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVFEEYKDYF 1256

Qy     113 VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
             | || ::::|   :  |||:|:||| |:
Db    1257 CISKWYLKEVKRKPYNLKLGCNWQFIPKDE 1286


RESULT 7
TETX_CLOTE
ID   TETX_CLOTE     STANDARD;      PRT;  1314 AA.
AC   P04958;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin).
OS   Clostridium tetani.
OG   Plasmid.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87053814; PubMed=3536478;
RA   Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA   Weller U., Hudel M., Habermann E., Niemann H.;
RT   "Tetanus toxin: primary structure, expression in E. coli, and
RT   homology with botulinum toxins.";
RL   EMBO J. 5:2495-2502(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CN3911;
RX   MEDLINE=87040747; PubMed=3774547;
RA   Fairweather N.F., Lyness V.A.;
RT   "The complete nucleotide sequence of tetanus toxin.";
RL   Nucleic Acids Res. 14:7809-7812(1986).
RN   [3]
RP   SEQUENCE OF 742-1314 FROM N.A.
RX   MEDLINE=86085672; PubMed=3510187;
RA   Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT   "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT   fragment C in Escherichia coli.";
RL   J. Bacteriol. 165:21-27(1986).
RN   [4]
RP   PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX   MEDLINE=90201034; PubMed=2108021;
RA   Krieglstein K., Henschen A., Weller U., Habermann E.;
RT   "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT   in tetanus toxin.";
RL   Eur. J. Biochem. 188:39-45(1990).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=92037649; PubMed=1935979;
RA   Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT   "Limited proteolysis of tetanus toxin. Relation to activity and
RT   identification of cleavage sites.";
RL   Eur. J. Biochem. 202:41-51(1991).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93010948; PubMed=1396558;
RA   Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA   Montecucco C.;
RT   "Tetanus toxin is a zinc protein and its inhibition of
RT   neurotransmitter release and protease activity depend on zinc.";
RL   EMBO J. 11:3577-3583(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX   MEDLINE=97475217; PubMed=9334741;
RA   Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA   Sax M.;
RT   "Structure of the receptor binding fragment HC of tetanus
RT   neurotoxin.";
RL   Nat. Struct. Biol. 4:788-792(1997).
CC   -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC       BOND OF SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
CC       SYNAPTOBREVIN.
CC   -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC       YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC       AND ARE NON-TOXIC AFTER SEPARATION.
CC   -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC       GANGLIOSIDE RECEPTORS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04436; CAA28033.1; -.
DR   EMBL; M12739; AAA23282.1; -.
DR   EMBL; X06214; CAA29564.1; -.
DR   PIR; A25689; BTCLTN.
DR   PDB; 1AF9; 29-APR-98.
DR   PDB; 1A8D; 14-OCT-98.
DR   MEROPS; M27.001; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
```

```
KW   3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    456       TETANUS TOXIN LIGHT CHAIN.
FT   CHAIN       457   1314       TETANUS TOXIN HEAVY CHAIN.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    233    233       BY SIMILARITY.
FT   METAL       236    236       ZINC (CATALYTIC) (BY SIMILARITY).
FT   TRANSMEM    226    246       POTENTIAL.
FT   TRANSMEM    669    689       POTENTIAL.
FT   DISULFID    438    466       INTERCHAIN.
FT   DISULFID   1076   1092
SQ   SEQUENCE   1314 AA;  150550 MW;  134C3657133EF81D CRC64;

  Query Match             22.4%;  Score 163;  DB 1;  Length 1314;
  Best Local Similarity   26.9%;  Pred. No. 2.9e-07;
  Matches   43;  Conservative   26;  Mismatches   49;  Indels   42;  Gaps    6;

Qy       5 LYRGTKFIIKKYASGNK-DNIVRNNDRVYINVVVKNKEY------------------RLA 45
           || | |||||:|   |: |: |:: | : : |   | |:                  |:
Db    1168 LYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVG 1227

Qy      46 TNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFH- 104
             ||    : | : |::: |:                |   :: | |:    :| :| |
Db    1228 YNAPGIPLYKKMEAVKLRDLK---------------TYSVQLKLYDDKNASLGLVGTHN 1271

Qy     105 -QFNNIAK---LVASNWYNRQIERSSRTLGCSWEFIPVDD 140
            | |       |:|||||   ::   : ||| | |:| |:
Db    1272 GQIGNDPNRDILIASNWYFNHLK--DKILGCDWYFVPTDE 1309


RESULT 8
BXG_CLOBO
ID   BXG_CLOBO      STANDARD;      PRT;  1296 AA.
AC   Q60393;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
DE   (Bontoxilysin G).
GN   BOTG.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=113 / 30;
RX   MEDLINE=94092745; PubMed=8268233;
RA   Campbell K., Collins M.D., East A.K.;
RT   "Nucleotide sequence of the gene coding for Clostridium botulinum
RT   (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT   with other clostridial neurotoxins.";
RL   Biochim. Biophys. Acta 1216:487-491(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74162; CAA52275.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    441       BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT   CHAIN       442   1296       BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    435    449       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1296 AA;  149013 MW;  DC8E47E15F665C31 CRC64;

  Query Match             21.8%;  Score 158.5;  DB 1;  Length 1296;
  Best Local Similarity   29.9%;  Pred. No. 7.4e-07;
  Matches   46;  Conservative   28;  Mismatches   59;  Indels   21;  Gaps    7;

Qy       4 SLYRGTKFIIKKYASG---NKDNIVRNNDRVYINV-VVKNKEYRLATNASQAGVEKILSA 59
           :|| | :|||||  ::    | |||||  | :|:|:  : :: ||:    :   ::  |
Db    1143 NLYLGLRFIIKKASNSRNINNDNIVREGDYIYLNIDNISDESYRVYVVLVNSKEIQTQLFL 1202

Qy      60 LEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQF-----------NN 108
             | |      |: :| |  :  |  |:: | : :    |  |  :|           :|
Db    1203 APINDDPTFYDVLQIK-KYYEKTTYNCQI-LCEKDTKTFGLFGIGKFVKDYGYVWDTYDN 1260

Qy     109 IAKLVASNWYNRQIER--SSRTLGCSWEFIPVDD 140
             | || |:|    :   |||:|:|||||:
Db    1261 Y--FCISQWYLRRISENINKLRLGCNWQFIPVDE 1292


RESULT 9
BXC1_CLOBO
ID   BXC1_CLOBO     STANDARD;      PRT;  1290 AA.
AC   P18640;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE   (Bontoxilysin C1).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90370487; PubMed=2204031;
RA   Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
RA   Boquet P., Popoff M.R.;
RT   "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL   Nucleic Acids Res. 18:4924-4924(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=91024998; PubMed=2222445;
RA   Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA   Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT   "The complete nucleotide sequence of the gene coding for botulinum
RT   type C1 toxin in the C-ST phage genome.";
RL   Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN   [3]
RP   SEQUENCE OF 2-25.
```

```
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=88153072; PubMed=2450068;
RA   Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
RA   Oguma K.;
RT   "Establishment of a monoclonal antibody recognizing an antigenic site
RT   common to Clostridium botulinum type B, C1, D, and E toxins and
RT   tetanus toxin.";
RL   Infect. Immun. 56:898-902(1988).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94038966; PubMed=7901002;
RA   Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT   "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
RT   cleaving HPC-1/syntaxin.";
RL   EMBO J. 12:4821-4828(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
CC       STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC       BACTERIOPHAGE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66433; CAA47060.1; -.
DR   EMBL; X72793; CAA51313.1; -.
DR   EMBL; X53751; CAA37780.1; -.
DR   EMBL; D90210; BAA14235.1; -.
DR   EMBL; X62389; CAA44263.1; -.
DR   PIR; S11291; S11291.
DR   PIR; A35396; A35396.
DR   PIR; A43503; A43503.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    448       BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
FT   CHAIN       449   1290       BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
FT   METAL       228    228       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    229    229       BY SIMILARITY.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    452       INTERCHAIN (PROBABLE).
FT   CONFLICT     84     84       P -> T (IN REF. 2).
SQ   SEQUENCE   1290 AA;  148734 MW;  71FBE379F97129E8 CRC64;

  Query Match             13.8%;  Score 100;  DB 1;  Length 1290;
  Best Local Similarity   23.6%;  Pred. No. 0.18;
  Matches   37;  Conservative   21;  Mismatches   77;  Indels   22;  Gaps    4;

Qy       2 NSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRL--------ATNASQAGV 53
           |:    | | |||:      |  ||  | :| :: : || | |        | | |   :
Db    1138 NNDFNEGYKIIIKRIRGNTNDTRVRGGDILYFDMTINNKAYNLFMKNETMYADNHSTEDI 1197

Qy      54 EKILSALEIPDVGNLSQVVVMKSKNDQGITNKC-KMNLQDNNGNDIGFIGFHQF------ 106
             |    :  |: :     :    |     ::  | |    | : |  || ::|
Db    1198 YAIGLREQTKDINDNIIFQIQPMNNTYYYASQIFKSNFNGENISGICSIGTYRFRLGGDW 1257

Qy     107 ---NNIAKLVASNWYNRQIERSSRTLGCSWEFIPVDD 140
             | :   |    |   :| :|      | |:|| :
Db    1258 YRHNYLVPTVKQGNYASLLESTS----THWGFVPVSE 1290


RESULT 10
MURI_FUSNN
ID   MURI_FUSNN     STANDARD;      PRT;   264 AA.
AC   Q8REE6;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutamate racemase (EC 5.1.1.3).
GN   MURI OR FN1161.
OS   Fusobacterium nucleatum (subsp. nucleatum).
OC   Bacteria; Fusobacteria; Fusobacterium.
OX   NCBI_TaxID=76856;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25586;
RX   MEDLINE=21886394; PubMed=11889109;
RA   Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA   Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA   Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA   Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA   Fonstein M., Kyrpides N., Overbeek R.;
RT   "Genome sequence and analysis of the oral bacterium Fusobacterium
RT   nucleatum strain ATCC 25586.";
RL   J. Bacteriol. 184:2005-2018(2002).
CC   -!- FUNCTION: Provides the (R)-glutamic acid required for cell wall
CC       biosynthesis (By similarity).
CC   -!- CATALYTIC ACTIVITY: L-glutamate = D-glutamate.
CC   -!- PATHWAY: Peptidoglycan biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE ASPARTATE/GLUTAMATE RACEMASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE010622; AAL95357.1; -.
DR   PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR   PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
KW   Peptidoglycan synthesis; Cell wall; Isomerase; Complete proteome.
SQ   SEQUENCE   264 AA;  29882 MW;  F69A0C16BEAB6AF9 CRC64;

  Query Match             12.2%;  Score 88.5;  DB 1;  Length 264;
  Best Local Similarity   26.7%;  Pred. No. 0.36;
  Matches   39;  Conservative   22;  Mismatches   54;  Indels   31;  Gaps    6;

Qy      16 YASGNKDNIVRNNDRVYINVVVKN--KEYRLATN-ASQAGVEKILSALEIPDVGNLSQVV 72
           | ||   | ::      ::  |||  |   :| | || | :: :     :| :| :   |
Db      44 YGSGKTKNELQKLTERILDFFVKNNCKLVIVACNTASTAAIDYLREKFSLPIIGIIEAGV 103

Qy      73 VMKSKN----------------DQGITNKCKMNLQDNNGNDIGFIGFHQ--------FNN 108
             : |||                 |  || ||   : |  :|  | | |        |:|
Db     104 KIASKNTKNKNIAVISTKFTAESHGYKNKAKMLDSELNVKEIACIEFAQMIETGWDTFDN 163

Qy     109 IAKLVASNWYNRQIERSSRT--LGCS 132
             :|:  | |  :| ::: |  |||:
```

```
Db    164 RKELL--NKYLSEIPKNADTLVLGCT 187


RESULT 11
YCF1_ARATH
ID   YCF1_ARATH     STANDARD;      PRT;  1786 AA.
AC   P56785;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 213.7 kDa protein ycf1.
GN   YCF1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20039611; PubMed=10574454;
RA   Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;
RT   "Complete structure of the chloroplast genome of Arabidopsis
RT   thaliana.";
RL   DNA Res. 6:283-290(1999).
CC   -!- FUNCTION: NOT YET KNOWN.
CC   -!- SIMILARITY: BELONGS TO THE YCF1 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AP000423; BAA84445.1; -.
DR   EMBL; AP000423; BAA84433.1; -.
KW   Chloroplast; Hypothetical protein.
SQ   SEQUENCE   1786 AA;  213727 MW;  CFFD2A4D776D7E5D CRC64;

  Query Match             10.9%;  Score 79.5;  DB 1;  Length 1786;
  Best Local Similarity   25.5%;  Pred. No. 20;
  Matches   40;  Conservative   26;  Mismatches   58;  Indels   33;  Gaps    8;

Qy      6 YRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKN-------KEYRLATNASQAGVEKILS 58
          :  || |: ||   |::|  :  : :|    :||         | |  : ||| |   ||
Db   1103 FESTKKILDKYIYKNEENGEKKKNTLYFISTIKNLISNKKKMSYDLC-SLSQAYVFYKLS 1161

Qy     59 ALEIPDVGNLSQVV---------VMKSK-----NDQGI-----TNKCKMNLQDNNGNDIG 99
           ::: :   |  |:          :|:|      : ||      ||  :| : |   :
Db   1162 QIKVSNFCKLKAVLEYNICITSFFVKNKIKVFFQEHGIFHYELKNKTFLNSEVNQWKN-W 1220

Qy    100 FIGFHQFN----NIAKLVASNWYNRQIERSSRTLGCS 132
            :|:|    : |:||  || |: | : |  |  |
Db   1221 LRSQYQYNLPQISWARLVTQNWKNK-INKDSLVLNPS 1256


RESULT 12
Y008_HUMAN
ID   Y008_HUMAN     STANDARD;      PRT;   765 AA.
AC   Q15398;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein KIAA0008.
GN   KIAA0008.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=96051387; PubMed=7584026;
RA   Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
RA   Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT   "Prediction of the coding sequences of unidentified human genes. I.
RT   The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT   analysis of randomly sampled cDNA clones from human immature myeloid
RT   cell line KG-1.";
RL   DNA Res. 1:27-35(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye, and Lung;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D13633; BAA02797.1; -.
DR   EMBL; BC010658; AAH10658.1; -.
DR   EMBL; BC016276; AAH16276.1; -.
DR   InterPro; IPR005026; GKAP.
DR   Pfam; PF03359; GKAP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   765 AA;  85668 MW;  00AFF91A02387EA1 CRC64;

  Query Match             10.9%;  Score 79;  DB 1;  Length 765;
  Best Local Similarity   21.9%;  Pred. No. 8.8;
  Matches   32;  Conservative   27;  Mismatches   39;  Indels   48;  Gaps    8;

Qy     20 NKDNIVRNNDRVYINV---------VVKNKEYRLA---------TNASQAGVEKILS--- 58
          ::::::  |:    |:          ::  | |:|          |  |:  ||: |
Db    309 HEEHVLNKNEATTKNLNGLPIKEVPSLERNEGRIAQPHHGVPYFRNILQSETEKLTSHCF 368

Qy     59 ------ALEIPD---------VGNLSQVVVMKSKNDQGITNKC--KMNLQDNNGNDI-GF 100
                |:|||          ||    ::  : |  :|: : |  |  :::    |: ||
Db    369 EWDRKLELDIPDDAKDLIRTAVGQTRLLMKERFKQFEGLVDDCEYKRGIKETTCTDLDGF 428

Qy    101 ---IGF------HQFNNIAKLVASNW 117
              : |      |:|||: ||  | |
Db    429 WDMVSFQIEDVIHKFNNLIKLEESGW 454


RESULT 13
WZC_ECOLI
ID   WZC_ECOLI      STANDARD;      PRT;   720 AA.
AC   P76387; P71236; O08003; O08004;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase wzc (EC 2.7.1.112).
GN   WZC OR B2060.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=96326333; PubMed=8759852;
RA   Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;
RT   "Organization of the Escherichia coli K-12 gene cluster responsible
RT   for production of the extracellular polysaccharide colanic acid.";
RL   J. Bacteriol. 178:4885-4893(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
```

```
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [4]
RP   CHARACTERIZATION.
RC   STRAIN=K12 / JM109;
RX   MEDLINE=99287830; PubMed=10348860;
RA   Vincent C., Doublet P., Grangeasse C., Vaganay E., Cozzone A.J.,
RA   Duclos B.;
RT   "Cells of Escherichia coli contain a protein-tyrosine kinase, Wzc, and
RT   a phosphotyrosine-protein phosphatase, Wzb.";
RL   J. Bacteriol. 181:3472-3477(1999).
RN   [5]
RP   CHARACTERIZATION.
RC   STRAIN=K12 / JM109;
RX   MEDLINE=20545593; PubMed=11090276;
RA   Vincent C., Duclos B., Grangeasse C., Vaganay E., Riberty M.,
RA   Cozzone A.J., Doublet P.;
RT   "Relationship between exopolysaccharide production and protein-
RT   tyrosine phosphorylation in gram-negative bacteria.";
RL   J. Mol. Biol. 304:311-321(2000).
RN   [6]
RP   CHARACTERIZATION, AND MUTAGENESIS.
RC   STRAIN=K12 / JM109;
RX   MEDLINE=21850696; PubMed=11751920;
RA   Grangeasse C., Doublet P., Cozzone A.J.;
RT   "Tyrosine phosphorylation of protein kinase Wzc from Escherichia coli
RT   K12 occurs through a two-step process.";
RL   J. Biol. Chem. 277:7127-7135(2002).
CC   -!- FUNCTION: Required for the extracellular polysaccharide colanic
CC       acid synthesis. The autophosphorylated form is inactive. Probably
CC       involved in the export of colanic acid from the cell to medium.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- ENZYME REGULATION: Dephosphorylated and activated by wzb.
CC   -!- PATHWAY: Exopolysaccharide biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Inner membrane (Probable).
CC   -!- PTM: Autophosphorylated. Seems to be phosphorylated through a
CC       cooperative two-step mechanism. First, Tyr-569 is phosphorylated
CC       in an intramolecular reaction that generates a significant
CC       increase of protein kinase activity. Then Tyr-708, Tyr-710, Tyr-
CC       711, Tyr-713 and Tyr-715 are phosphorylated in an intermolecular
CC       Tyr-569-dependent reaction.
CC   -!- MISCELLANEOUS: Additional site-directed mutagenesis experiments
CC       indicated that the tyrosine residues at positions 708, 710, 711,
CC       713 and 715 are phosphorylation sites, whereas tyrosine at
CC       position 705 is not.
CC   -!- SIMILARITY: BELONGS TO THE ETK/WZC FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U38473; AAC77835.1; -.
DR   EMBL; AE000296; AAC75121.1; ALT_INIT.
DR   EMBL; D90843; BAA15913.1; -.
DR   EMBL; D90844; BAA15915.1; -.
DR   PhosSite; P76387; -.
DR   EcoGene; EG13568; wzc.
KW   Transferase; Tyrosine-protein kinase; Phosphorylation;
KW   Exopolysaccharide synthesis; Transmembrane; Inner membrane;
KW   Complete proteome.
FT   DOMAIN        1     31       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     32     52       POTENTIAL.
FT   DOMAIN       53    424       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    425    445       POTENTIAL.
FT   DOMAIN      446    720       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES     569    569       PHOSPHORYLATION (AUTO-) (PROBABLE).
FT   MOD_RES     708    708       PHOSPHORYLATION (PROBABLE).
FT   MOD_RES     710    710       PHOSPHORYLATION (PROBABLE).
FT   MOD_RES     711    711       PHOSPHORYLATION (PROBABLE).
FT   MOD_RES     713    713       PHOSPHORYLATION (PROBABLE).
FT   MOD_RES     715    715       PHOSPHORYLATION (PROBABLE).
FT   MUTAGEN     467    467       Y->F: NO LOSS OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     491    491       Y->F: NO LOSS OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     540    540       K->M: LOSS OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     569    569       Y->F: LOSS OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     636    636       Y->F: NO LOSS OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     668    668       Y->F: NO LOSS OF AUTOPHOSPHORYLATION.
SQ   SEQUENCE   720 AA;  79343 MW;  8F61017B5ECFFB45 CRC64;

  Query Match             10.3%;  Score 75;  DB 1;  Length 720;
  Best Local Similarity   25.6%;  Pred. No. 19;
  Matches   32;  Conservative   20;  Mismatches   55;  Indels   18;  Gaps    5;

Qy      29 DRVYINVVVKNKEYRLATNA-----SQAG----VEKILSALEIPDVGNLSQVVVMKSKND 79
           |:|:   |: || | |:::       |||     | :   :|       |:  | |  :
Db     149 DQVFTLNVLDNKNYTLSSDGGFSARGQAGQMLKKEGVTLMVEAIHASPGSEFTVTK-YST 207

Qy      80 QGITNKCKMNLQ-DNNGNDIGFIGF-------HQFNNIAKLVASNWYNRQIERSSRTLGC 131
            |: |:|: :|    || | | :          |  :|   :| |:  : ||| |
Db     208 LGMINQLQNSLTVTENGKDAGVLSLTYTGEDREQIRDILNSIARNYQEQNIERKSAEASK 267

Qy     132 SWEFI 136
           |  |:
Db     268 SLAFL 272


RESULT 14
BXD_CLOBO
ID   BXD_CLOBO      STANDARD;      PRT;  1276 AA.
AC   P19321;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
DE   (Bontoxilysin D).
GN   BOTD.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BVD/-3;
RX   MEDLINE=91016853; PubMed=2216736;
RA   Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
RA   Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
RT   "Nucleotide sequence of the gene encoding Clostridium botulinum
RT   neurotoxin type D.";
RL   Nucleic Acids Res. 18:5556-5556(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CB16;
RX   MEDLINE=93042276; PubMed=1420572;
```

```
RA   Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
RT   "The complete amino acid sequence of the Clostridium botulinum type D
RT   neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT   phage d-16 phi genome.";
RL   J. Vet. Med. Sci. 54:905-913(1992).
RN   [3]
RP   PARTIAL SEQUENCE.
RC   STRAIN=D-SA, and D-1873;
RX   MEDLINE=89339741; PubMed=2668193;
RA   Moriishi K., Syuto B., Kubo S., Oguma K.;
RT   "Molecular diversity of neurotoxins from Clostridium botulinum type D
RT   strains.";
RL   Infect. Immun. 57:2886-2891(1989).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-|-LEU-61 BOND OF
CC       SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
CC       STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC       BACTERIOPHAGE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54254; CAA38175.1; -.
DR   EMBL; S49407; AAB24244.1; -.
DR   PIR; S11455; S11455.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    442       BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
FT   CHAIN       443   1276       BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    437    450       INTERCHAIN (PROBABLE).
FT   VARIANT      15     16       ND -> PV (IN STRAIN D-SA).
FT   VARIANT      17     18       ND -> LQ (IN STRAIN D-1873).
FT   VARIANT     452    452       K -> Q (IN STRAIN D-SA).
FT   VARIANT     457    457       R -> T (IN STRAIN D-SA).
FT   VARIANT     457    457       R -> F (IN STRAIN D-1873).
FT   VARIANT     462    462       A -> D (IN STRAIN D-1873).
FT   VARIANT     489    489       K -> N (IN STRAIN CB16).
FT   VARIANT     644    644       N -> K (IN STRAIN CB16).
FT   VARIANT    1122   1122       Q -> R (IN STRAIN CB16).
SQ   SEQUENCE   1276 AA;  146871 MW;  C1EC50F46C8233E2 CRC64;

  Query Match             10.2%;  Score 74;  DB 1;  Length 1276;
  Best Local Similarity   21.9%;  Pred. No. 45;
  Matches   30;  Conservative   18;  Mismatches   39;  Indels   50;  Gaps    5;

Qy       3 SSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATN-----ASQAGVEKIL 57
           | || |    ||  :  |  : : | | : ::::  :::| :  :     |:| |
Db    1127 SKLYTGNPITIKSVSDKNPYSRILNGDNIILHMLYNSRKYMIIRDTDTIYATQGG----- 1181

Qy      58 SALEIPDVGNLSQVVVMKSKNDQGITNKCKMN------LQDNNGN-DIGFIGFHQFNNIA 110
                                    :|  |      || | ||  ||       :|
Db    1182 -------------------------ECSQNCVYALKLQSNLGNYGIGIF------SIK 1208

Qy     111 KLVASNWYNRQIERSSR 127
           :|: | |  ||  | |
Db    1209 NIVSKNKYCSQIFSSFR 1225


RESULT 15
NISP_LACLA
ID   NISP_LACLA     STANDARD;      PRT;   682 AA.
AC   Q07596;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Nisin leader peptide processing serine protease nisP precursor
DE   (EC 3.4.21.-).
GN   NISP.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIZO R5;
RX   MEDLINE=93239683; PubMed=8478324;
RA   van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,
RA   Kuipers O.P., de Vos W.M.;
RT   "Characterization of the Lactococcus lactis nisin A operon genes
RT   nisP, encoding a subtilisin-like serine protease involved in
RT   precursor processing, and nisR, encoding a regulatory protein
RT   involved in nisin biosynthesis.";
RL   J. Bacteriol. 175:2578-2588(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=6F3;
RX   MEDLINE=94213458; PubMed=8161176;
RA   Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,
RA   Hammelmann M., Entian K.-D.;
RT   "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
RT   6F3.";
RL   Appl. Environ. Microbiol. 60:814-825(1994).
RN   [3]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=95357326; PubMed=7630881;
RA   Siezen R.J., Rollema H.S., Kuipers O.P., de Vos W.M.;
RT   "Homology modelling of the Lactococcus lactis leader peptidase NisP
RT   and its interaction with the precursor of the lantibiotic nisin.";
RL   Protein Eng. 8:117-125(1995).
CC   -!- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC   -!- PATHWAY: Nisin biosynthesis; last step.
CC   -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC       an amide bond (Potential).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L11061; AAA25200.1; -.
DR   EMBL; X76884; CAA54210.1; -.
DR   HSSP; P29600; 1GCI.
DR   MEROPS; S08.059; -.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR000209; Peptidase_S8.
DR   Pfam; PF00082; Peptidase_S8; 1.
DR   PRINTS; PR00723; SUBTILISIN.
DR   PROSITE; PS00136; SUBTILASE_ASP; 1.
DR   PROSITE; PS00137; SUBTILASE_HIS; 1.
DR   PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
DR   PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
KW   Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   PROPEP       23    195       POTENTIAL.
FT   CHAIN       196    655       NISIN LEADER PEPTIDE PROCESSING SERINE
FT                                PROTEASE NISP.
FT   PROPEP       656    682       REMOVED BY SORTASE (POTENTIAL).
FT   ACT_SITE     259    259       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     306    306       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE     512    512       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   SITE         652    656       LPXTG SORTING SIGNAL (POTENTIAL).
FT   MOD_RES      655    655       AMIDE-LINKED TO CELL WALL (POTENTIAL).
FT   CONFLICT     500    500       A -> T (IN REF. 2).
SQ   SEQUENCE   682 AA;  74767 MW;  D5F29313F2983EC9 CRC64;

  Query Match             10.1%;  Score 73.5;  DB 1;  Length 682;
  Best Local Similarity   25.2%;  Pred. No. 25;
  Matches   37;  Conservative   22;  Mismatches   51;  Indels   37;  Gaps    8;

Qy      12 IIKKYASGNKDNIVR---------NNDRVYINVV----VKNKEYRLATNASQAGVEKILS 58
           :: ||   | : : |         | :|| :|:|     ||| : | |:  |
Db     525 VVDKYGIKNPNQLKRFLLMNSPEVNGNRV-LNIVDLLNGKNKAFSLDTDKGQD------D 577

Qy      59 ALEIPDVGNLSQV-VVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
           |:    : || :    || : |: |      |    |           |:||:| | |
Db     578 AINHKSMENLKESRDTMKQEQDKEIQRNTNNNFSIKN----------DFHNISKEVISVD 627

Qy     118 Y--NRQIERSSRTLGC----SWEFIPV 138
           |  |:::  :  : |     | | :||
Db     628 YNINQKMANNRNSRGAVSVRSQEILPV 654


Search completed: March 13, 2003, 11:40:38
Job time : 8.45533 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:47 ; Search time 20.9798 Seconds
(without alignments)
1374.968 Million cell updates/sec

Title: US-09-917-791-22
Perfect score: 727
Sequence: 1 LNSSLYRGTKFIIKKYASGN.....QIERSSRTLGCSWEFIPVDD 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	341	46.9	72	2	Q9R540	Q9r540 clostridium
2	282.5	38.9	1268	2	Q45851	Q45851 clostridium
3	277.5	38.2	1278	2	Q57236	Q57236 clostridium
4	260.5	35.8	1280	2	Q9ZAJ5	Q9zaj5 clostridium
5	205	28.2	1251	2	Q9K395	Q9k395 clostridium
6	197.5	27.2	1255	2	Q9FAR6	Q9far6 clostridium
7	169.5	23.3	1291	2	Q08077	Q08077 clostridium
8	162	22.3	451	2	Q9LA13	Q91a13 clostridium
9	161.5	22.2	441	2	Q9X708	Q9x708 clostridium
10	160.5	22.1	1291	2	Q93G71	Q93g71 clostridium
11	160.5	22.1	1291	2	Q933K0	Q933k0 clostridium
12	160.5	22.1	1291	2	Q9ZAJ8	Q9zaj8 clostridium
13	160	22.0	1310	2	Q93N27	Q93n27 clostridium
14	110	15.1	315	16	Q97FU1	Q97fu1 clostridium
15	100	13.8	1291	2	Q93HT3	Q93ht3 clostridium
16	95	13.1	1285	2	Q9LBR1	Q91br1 clostridium
17	95	13.1	1285	2	Q45967	Q45967 clostridium
18	88.5	12.2	264	16	Q8REE6	Q8ree6 fusobacteri
19	88	12.1	408	15	O11572	O11572 human immun
20	87.5	12.0	204	15	Q9DV80	Q9dv80 human immun
21	86.5	11.9	401	15	O11575	O11575 human immun
22	85.5	11.8	467	16	Q97N26	Q97n26 clostridium
23	85	11.7	1593	5	Q8T1T7	Q8t1t7 dictyosteli
24	84.5	11.6	848	15	Q9WC60	Q9wc60 human immun
25	83	11.4	245	16	Q8RDM8	Q8rdm8 fusobacteri
26	83	11.4	408	15	O11570	O11570 human immun
27	83	11.4	409	15	O11571	O11571 human immun
28	83	11.4	1483	3	Q96UL6	Q96ul6 diaporthe a
29	83	11.4	4550	5	O77336	O77336 plasmodium
30	82.5	11.3	479	10	Q9M6E4	Q9m6e4 nicotiana t
31	82	11.3	255	16	Q8XM40	Q8xm40 clostridium
32	82	11.3	408	15	O11567	O11567 human immun
33	81.5	11.2	235	2	Q93D91	Q93d91 streptococc
34	81.5	11.2	408	15	O11574	O11574 human immun
35	81.5	11.2	1077	1	Q977Q4	Q977q4 methanosarc
36	81.5	11.2	1869	12	Q8V2H2	Q8v2h2 camelpox vi
37	81.5	11.2	1869	12	Q8QPZ7	Q8qpz7 camelpox vi
38	81	11.1	336	12	Q91MU9	Q91mu9 lumpy skin
39	81	11.1	1301	5	Q8WSK5	Q8wsk5 plasmodium
40	81	11.1	1484	3	Q8TGV4	Q8tgv4 diaporthe
41	80	11.0	1122	3	Q8X1C2	Q8x1c2 phomopsis s
42	79.5	10.9	214	15	Q74058	Q74058 human immun
43	79.5	10.9	233	16	Q92CI5	Q92ci5 listeria in
44	79.5	10.9	307	16	Q8XKL0	Q8xkl0 clostridium
45	79.5	10.9	1807	5	Q8WSL1	Q8wsl1 plasmodium

ALIGNMENTS

RESULT 1
Q9R540
ID Q9R540 PRELIMINARY; PRT; 72 AA.
AC Q9R540;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Neurotoxin heavy chain 18 kDa fragment (Fragment).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE.
RX MEDLINE=94000342; PubMed=8397793;
RA Gimenez J.A., DasGupta B.R.;
RT "Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72,
RT 45, 42, and 18 kD fragments.";
RL J. Protein Chem. 12:351-363(1993).
DR HSSP; P10845; 3BTA.
SQ SEQUENCE 72 AA; 8165 MW; B7A959576A615E18 CRC64;

Query Match 46.9%; Score 341; DB 2; Length 72;
Best Local Similarity 98.6%; Pred. No. 1.7e-24;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 62

Qy      61 EIPDVGNLSQ 70
           |||||||| |
Db      63 EIPDVGNLYQ 72
```

RESULT 2
Q45851
ID Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Neurotoxin type F.
GN   BONT /F.
OS   Clostridium baratii.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1561;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93252228; PubMed=8486245;
RA   Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA   Richardson P.T.;
RT   "Nucleotide sequence of the gene coding for Clostridium barati type F
RT   neurotoxin: Comparison with other clostridial neurotoxins.";
RL   FEMS Microbiol. Lett. 108:175-182(1993).
DR   EMBL; X68262; CAA48329.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1268 AA;  145513 MW;  963040091AC15ED2 CRC64;

  Query Match             38.9%;  Score 282.5;  DB 2;  Length 1268;
  Best Local Similarity   45.3%;  Pred. No. 1.5e-17;
  Matches   63;  Conservative   19;  Mismatches   44;  Indels   13;  Gaps    5;

Qy       2 NSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILS 58
           |: || | : ||:|  |    | || || || ||||||  | ||:|  : | : ||| :
Db    1131 NARLYTGVEVIIRKVGSTDTSNTDNFVRKNDTVYINVVDGNSEYQLYADVSTSAVEKTIK 1190

Qy      59 ALEIPDVG-NLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
             | :   | :|:::| |     | : | || : |||||| :||| ||   ||||:|
Db    1191 LRRISNSNYNSNQMIIMDS-----IGDNCTMNFKTNNGNDIGLLGFH-LNN---LVASSW 1241

Qy     118 YNRQIERSSRTLGCSWEFI 136
           | : |  ::|  || | ||
Db    1242 YYKNIRNNTRNNGCFWSFI 1260


RESULT 3
Q57236
ID   Q57236      PRELIMINARY;      PRT;  1278 AA.
AC   Q57236; Q45863;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Botulinum neurotoxin type F (BONT/F protein).
GN   BONT/F.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 10281;
RA   Hutson R.A., Collins M.D.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 635-1000 FROM N.A.
RC   STRAIN=NCTC 1028;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-27 FROM N.A.
RC   STRAIN=LANGELAND;
RX   MEDLINE=98404102; PubMed=9732534;
RA   East A.K., Bhandari M., Hielm S., Collins M.D.;
RT   "Analysis of the botulinum neurotoxin type F gene clusters in
RT   proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT   barati.";
RL   Curr. Microbiol. 37:262-268(1998).
DR   EMBL; X81714; CAA57358.1; -.
DR   EMBL; L35496; AAA23210.1; -.
DR   EMBL; X70821; CAA50152.1; -.
DR   EMBL; X99064; CAA67512.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin.
SQ   SEQUENCE   1278 AA;  147073 MW;  A1BE1318431D6918 CRC64;

  Query Match             38.2%;  Score 277.5;  DB 2;  Length 1278;
  Best Local Similarity   44.6%;  Pred. No. 4.5e-17;
  Matches   62;  Conservative   19;  Mismatches   45;  Indels   13;  Gaps    4;

Qy       2 NSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILS 58
           |: || | : ||:|  |    | || || || ||||| :: ||||  : | |  |||:
Db    1140 NTRLYTGVEVIIRKNGSTDISNTDNFVRKNDLAYINVVDRDVEYRLYADISIAKPEKIIK 1199

Qy      59 ALEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
             :   :  | | |::|| |     | | | || |:||| :|| :|||  |    ||||:|
Db    1200 LIRTSNSNNSLGQIIVMDS-----IGNNCTMNFQNNNGGNIGLLGFHSNN----LVASSW 1250

Qy     118 YNRQIERSSRTLGCSWEFI 136
           |   | ::: : || | ||
Db    1251 YYNNIRKNTSSNGCFWSFI 1269


RESULT 4
Q9ZAJ5
ID   Q9ZAJ5      PRELIMINARY;      PRT;  1280 AA.
AC   Q9ZAJ5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13631; CAA73972.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
```

```
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1280 AA;  147487 MW;  D0F748976EBC222C CRC64;

  Query Match             35.8%;  Score 260.5;  DB 2;  Length 1280;
  Best Local Similarity   42.9%;  Pred. No. 1.7e-15;
  Matches   60;  Conservative   19;  Mismatches   48;  Indels   13;  Gaps    4;

Qy       1 LNSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKIL 57
           ||  || | : ||:| |     | || || ||  |||||    ||||  : |    |||:
Db    1142 LNYKLYEGVEVIIRKNAPIDISNTDNFVRKNDLAYINVVDHGVEYRLYADISITKSEKII 1201

Qy      58 SALEIPDVG-NLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASN 116
             :   :   :| |::|| |     | | | || |:|:|::|| :|||  :    ||||:
Db    1202 KLIRTSNPNDSLGQIIVMDS-----IGNNCTMNFQNNDGSNIGLLGFHSDD----LVASS 1252

Qy     117 WYNRQIERSSRTLGCSWEFI 136
           ||   | |:: : || | ||
Db    1253 WYYNHIRRNTSSNGCFWSFI 1272


RESULT 5
Q9K395
ID   Q9K395      PRELIMINARY;      PRT;  1251 AA.
AC   Q9K395;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Type E botulinum toxin.
GN   BONT/E.
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 095;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (LCL 095) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 155;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
RA   Kato H., Nakamura S., Karasawa T.;
RT   "C. butyricum (LCL 155) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1899;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1897;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1898;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1886;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1886) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1887;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1889;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1889) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1890;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1890) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KZ 1891;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (KZ 1891) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL 063;
RA   Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA   Karasawa T.;
RT   "C. butyricum (LCL 063) gene for type E botulinum toxin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037714; BAB03522.1; -.
DR   EMBL; AB037704; BAB03512.1; -.
DR   EMBL; AB037705; BAB03513.1; -.
DR   EMBL; AB037706; BAB03514.1; -.
DR   EMBL; AB037707; BAB03515.1; -.
DR   EMBL; AB037708; BAB03516.1; -.
DR   EMBL; AB037709; BAB03517.1; -.
DR   EMBL; AB037710; BAB03518.1; -.
DR   EMBL; AB037711; BAB03519.1; -.
DR   EMBL; AB037712; BAB03520.1; -.
DR   EMBL; AB037713; BAB03521.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1251 AA;  143751 MW;  2021F4E427070296 CRC64;

  Query Match             28.2%;  Score 205;  DB 2;  Length 1251;
  Best Local Similarity   36.2%;  Pred. No. 2.4e-10;
  Matches   50;  Conservative   20;  Mismatches   54;  Indels   14;  Gaps    5;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSA 59
             | : ||.| |  |::   |   |  || ||:|||| :  :  | |  : :    || : :
Db    1117 LANRLYSGIKVKIQRVNDSSTNDRFVRKNDQVYINYISNSSSYSLYADTNTTDKEKTIKS 1176

Qy      60 LEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWY 118
                ||  :|||| |     : | | || ::||||:|| :||        :||| ||
Db    1177 ---SSSGNRFNQVVVMNS-----VGNNCTMNFKNNNGNNIGLLGF----KADTVVASTWY 1224

Qy     119 NRQIERSSRTLGCSWEFI 136
                :   : : || | ||
Db    1225 YTHMRDHTNSNGCFWNFI 1242
```

```
RESULT 6
Q9FAR6
ID   Q9FAR6      PRELIMINARY;      PRT;  1255 AA.
AC   Q9FAR6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Type E botulinum toxin.
GN   BONT/E.
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BL 6340/ATCC 43755/BL 5520/KZ 147;
RX   MEDLINE=20509829; PubMed=11055954;
RA   Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y.,
RA   Yamakawa K., Oguma K., Sakaguchi Y., Nakamura S.;
RT   "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium
RT   butyricum Strains.";
RL   Appl. Environ. Microbiol. 66:4992-4997(2000).
DR   EMBL; AB039264; BAB12249.1; -.
DR   HSSP; P10845; 3BTA.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1255 AA;  143918 MW;  1B557B9D85CD8E4D CRC64;

  Query Match             27.2%;  Score 197.5;  DB 2;  Length 1255;
  Best Local Similarity   38.4%;  Pred. No. 1.2e-09;
  Matches   53;  Conservative   20;  Mismatches   52;  Indels   13;  Gaps    6;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSA 59
           | : || | |  |::   |   ||:|| ||:|||| |  :| : |   |  |   |
Db    1120 LANRLYSGIKVKIQRVNNSSTNDNLVRKNDQVYINFVA-SKTHLLPLYADTATTNK-EKT 1177

Qy      60 LEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWY 118
           ::|   ||  :|||| |     : | | || ::|||:|| :||        :||| ||
Db    1178 IKISSSGNRFNQVVVMNS-----VGNNCTMNFKNNNGNNIGLLGF----KADTVVASTWY 1228

Qy     119 NRQIERSSRTLGCSWEFI 136
             :  :: : |  | ||
Db    1229 YTHMRDNTNSNGFFWNFI 1246


RESULT 7
Q08077
ID   Q08077      PRELIMINARY;      PRT;  1291 AA.
AC   Q08077;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Botulinum neurotoxin type B (EC 3.4.24.-) (BONT/B).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EKLUND 17B ATCC25765;
RX   MEDLINE=94122659; PubMed=7764370;
RA   Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT   "Nucleotide sequence of the gene coding for non-proteolytic
RT   Clostridium botulinum type B neurotoxin: comparison with other
RT   clostridial neurotoxins.";
RL   Curr. Microbiol. 28:101-110(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES SYNAPTOBREVIN-2.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC       NEUROTOXIN.
CC   -!- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC       REGION.
DR   EMBL; X71343; CAA50482.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
SQ   SEQUENCE   1291 AA;  150513 MW;  71BCAFE23D69FAAA CRC64;

  Query Match             23.3%;  Score 169.5;  DB 2;  Length 1291;
  Best Local Similarity   28.3%;  Pred. No. 4.8e-07;
  Matches   43;  Conservative   39;  Mismatches   53;  Indels   17;  Gaps    7;

Qy       4 SLYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRL-ATNASQAGVEKILSAL 60
           :|| | ||||:: ::     |:|||  | :::::|: ::|:|: |    :   ||:  ::
Db    1138 NLYIGEKFIIRRESNSQSINDDIVRKEDYIHLDLVLHHEEWRVYAYKYFKEQEEKLFLSI 1197

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIAKL------ 112
           | |     : : :|  ::|  :  |::     : : :||| || |:|     |
Db    1198 -ISDSNEFYKTIEIKEYDEQP-SYSCQLLFKKDEESTDDIGLIGIHRFYESGVLRKKYKD 1255

Qy     113 --VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
               | || ::::|      |||:|:||| |:
Db    1256 YFCISKWYLKEVKRKPYKSNLGCNWQFIPKDE 1287


RESULT 8
Q9LA13
ID   Q9LA13      PRELIMINARY;      PRT;  451 AA.
AC   Q9LA13;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tetanus toxin (Fragment).
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=20886;
RA   He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT   "Fragment C of Tetanus Toxin.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF154828; AAF73267.1; -.
DR   HSSP; P04958; 1A8D.
DR   InterPro; IPR001064; Crystallin.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   451 AA;  51823 MW;  69A8C5F030E6CD8E CRC64;

  Query Match             22.3%;  Score 162;  DB 2;  Length 451;
  Best Local Similarity   27.5%;  Pred. No. 6.6e-07;
```

```
  Matches   44;  Conservative   25;  Mismatches   49;  Indels   42;  Gaps    6;

Qy       5 LYRGTKFIIKKYASGNK-DNIVRNNDRVYINVVVKNKEY-----------------RLA 45
           || | |||||:|   |: |: ||: | : : |   | |:                  |:
Db     305 LYSGLKFIIKRYTPNNEIDSFVRSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVG 364

Qy      46 TNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFH- 104
           ||    : | : |::: |:                 |   :: | |:    :| :| |
Db     365 YNAPGIPLYKKMEAVKLRDLK---------------TYSVQLKLYDDKDASLGLVGTHN 408

Qy     105 -QFNNIAK---LVASNWYNRQIERSSRTLGCSWEFIPVDD 140
            |  |      |:|||||   ::   :|| | | |:| |:
Db     409 GQIGNDPNRDILIASNWYFNHLK--DKTLTCDWYFVPTDE 446


RESULT 9
Q9X708
ID   Q9X708      PRELIMINARY;      PRT;   441 AA.
AC   Q9X708;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343691; PubMed=10413679;
RA   Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA   Schiavo G.;
RT   "Functional characterisation of tetanus and botulinum neurotoxins
RT   binding domains.";
RL   J. Cell Sci. 112:2715-2724(1999).
DR   EMBL; AJ242628; CAB43706.1; -.
DR   HSSP; P10845; 3BTA.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     441    441
SQ   SEQUENCE   441 AA;  52772 MW;  721D0B468E8C95A4 CRC64;

  Query Match             22.2%;  Score 161.5;  DB 2;  Length 441;
  Best Local Similarity   27.9%;  Pred. No. 7.2e-07;
  Matches   41;  Conservative   33;  Mismatches   58;  Indels   15;  Gaps    5;

Qy       4 SLYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALE 61
           :|| | ||||:: ::     |:|||  | :|::    |:|:|:          || |
Db     296 NLYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNSNREWRVYAYKDFKEEEKKLVLAN 355

Qy      62 IPDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIAKLV------ 113
           | |     : : :|  ::|  |  |::     : : ::|| || |:|     ::
Db     356 IYDSNEFYKTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVLKDYKNY 414

Qy     114 --ASNWYNRQIERS--SRTLGCSWEFI 136
             | || ::::|   :  |||:|:||
Db     415 FCISKWYLKEVKRKPYNPNLGCNWQFI 441


RESULT 10
Q93G71
ID   Q93G71      PRELIMINARY;      PRT;  1291 AA.
AC   Q93G71;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Neurotoxin type B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1436;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF295926; AAK97132.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1291 AA;  150824 MW;  D7CA07BAE2EB8CD2 CRC64;

  Query Match             22.1%;  Score 160.5;  DB 2;  Length 1291;
  Best Local Similarity   28.0%;  Pred. No. 3.3e-06;
  Matches   42;  Conservative   31;  Mismatches   62;  Indels   15;  Gaps    5;

Qy       5 LYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEI 62
           || | ||||:: ::     |:|||  | :|::    |:|:|:          |: |    |
Db    1139 LYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNLNQEWRVYIYKYFKKEEEKLFLAPI 1198

Qy      63 PDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIA--------KL 112
             |       : :|  ::|  |  |::     : : ::|| || |:|
Db    1199 SDSDEFYNTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVFKEYKDYF 1257

Qy     113 VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
             | || ::::|   :  |||:|:|||  |:
Db    1258 CISKWYLKEVKRKPYNSKLGCNWQFIPKDE 1287


RESULT 11
Q933K0
ID   Q933K0      PRELIMINARY;      PRT;  1291 AA.
AC   Q933K0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Type B cryptic neurotoxin.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=593, AND 588;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF300466; AAL11499.1; -.
DR   EMBL; AF300465; AAL11498.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW   Neurotoxin.
SQ   SEQUENCE   1291 AA;  150843 MW;  7AC1737B0FA5A151 CRC64;

  Query Match             22.1%;  Score 160.5;  DB 2;  Length 1291;
  Best Local Similarity   28.0%;  Pred. No. 3.3e-06;
  Matches   42;  Conservative   31;  Mismatches   62;  Indels   15;  Gaps    5;

Qy       5 LYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEI 62
           || | ||||:: ::     |:|||  | :|::    |:|:|:          |: |    |
Db    1139 LYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNLNQEWRVYMYKYFKKEEEKLFLAPI 1198

Qy      63 PDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIA--------KL 112
             |       : :|  ::|  |  |::     : : ::|| || |:|
Db    1199 SDSDEFYNTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVFKEYKDYF 1257

Qy     113 VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
```

```
           | || ::::|   :  |||:|:||| |:
Db   1258 CISKWYLKEVKRKPYNSKLGCNWQFIPKDE 1287


RESULT 12
Q9ZAJ8
ID   Q9ZAJ8      PRELIMINARY;      PRT;  1291 AA.
AC   Q9ZAJ8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13630; CAA73968.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1291 AA;  150840 MW;  E4D3B0E46AB2E735 CRC64;

  Query Match             22.1%;  Score 160.5;  DB 2;  Length 1291;
  Best Local Similarity   28.0%;  Pred. No. 3.3e-06;
  Matches   42;  Conservative   31;  Mismatches   62;  Indels   15;  Gaps    5;

Qy       5 LYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEI 62
           || | ||||:: ::     |:||| | :|:: |:|:|:         |: |    |
Db    1139 LYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNLNQEWRVYMYKYFKKEEEKLFLAPI 1198

Qy      63 PDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIA--------KL 112
           |       : :|  ::|  |  |::     : : ::|| || |:|
Db    1199 SDSDEFYNTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVFKEYKDYF 1257

Qy     113 VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
           | || ::::|   :  |||:|:||| |:
Db    1258 CISKWYLKEVKRKPYNSKLGCNWQFIPKDE 1287


RESULT 13
Q93N27
ID   Q93N27      PRELIMINARY;      PRT;  1310 AA.
AC   Q93N27;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Tetanus toxin (Fragment).
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shumin Z., Dianliang L.;
RT   "Cloning and sequence analysis of tetanus toxin gene.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389424; AAK72964.2; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT   NON_TER       1       1
FT   NON_TER    1310    1310
SQ   SEQUENCE   1310 AA;  150316 MW;  9EADDC914418E450 CRC64;

  Query Match             22.0%;  Score 160;  DB 2;  Length 1310;
  Best Local Similarity   27.7%;  Pred. No. 3.7e-06;
  Matches   44;  Conservative   24;  Mismatches   49;  Indels   42;  Gaps    6;

Qy       5 LYRGTKFIIKKYASGNK-DNIVRNNDRVYINVVVKNKEY-----------------RLA 45
           || | |||||:|   |: |: ||: | : : |   | |:                 |:
Db    1170 LYSGLKFIIKRYTPNNEIDSFVRSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVG 1229

Qy      46 TNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFH- 104
             ||    : | : |::: |:              |   :: | |:    :| :| |
Db    1230 YNAPGIPLYKKMEAVKLRDLK---------------TYSVQLKLYDDKDASLGLVGTHN 1273

Qy     105 -QFNNIAK---LVASNWYNRQIERSSRTLGCSWEFIPVD 139
            |  |      |:|||||   ::   :|| | | |:| |
Db    1274 GQIGNDPNRDILIASNWYFNHLK--DKTLTCDWYFVPTD 1310


RESULT 14
Q97FU1
ID   Q97FU1      PRELIMINARY;      PRT;   315 AA.
AC   Q97FU1;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Hypothetical protein CAC2635.
GN   CAC2635.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007761; AAK80582.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   315 AA;  36592 MW;  1188CD3EE3A22124 CRC64;

  Query Match             15.1%;  Score 110;  DB 16;  Length 315;
  Best Local Similarity   32.6%;  Pred. No. 0.028;
  Matches   44;  Conservative   21;  Mismatches   52;  Indels   18;  Gaps    8;

Qy      12 IIKKY---ASGNKDNIV--RNNDRVYINVVVKNKEYRL--ATNASQAGVEKILSALE--I 62
           |||||     | ||||     ||::  | |  |:| | |  :   :  |  |  | |:  :
Db      47 IIKKYDVTGDGLKDNIYIKINNNKYSIAVKSKDKIYTLQPSHKLNSLGTYKSYSPLKLIL 106

Qy      63 PDVG--NLSQVVVMKSKNDQG-----ITNKCKM-NLQDNNGNDIGFIGFHQFNNIAKLVA 114
           ||   |:|::    |:|:       | :| |  ::     | ||||  |  |   |:::
Db     107 TDVSRDNISELFTQSSENNTNLQHLFIWDKTKFKDILHTTNNIIGFIDVHN-NKTPKIIS 165

Qy     115 SNWYNRQIERSSRTL 129
           ||: |  :  |:  |
Db     166 SNFNNNDLVFSNYIL 180


RESULT 15
```

```
Q93HT3
ID   Q93HT3      PRELIMINARY;      PRT;  1291 AA.
AC   Q93HT3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Neurotoxin.
GN   NT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-YOICHI;
RX   MEDLINE=21534265; PubMed=11676492;
RA   Sagane Y., Kouguchi H., Watanabe T., Sunagawa H., Inoue K.,
RA   Fujinaga Y., Oguma K., Ohyama T.;
RT   "Role of C-Terminal Region of HA-33 Component of Botulinum Toxin in
RT   Hemagglutination.";
RL   Biochem. Biophys. Res. Commun. 288:650-657(2001).
DR   EMBL; AB061780; BAB71749.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ   SEQUENCE   1291 AA;  148869 MW;  4A21DB35B8743CF8 CRC64;

  Query Match             13.8%;  Score 100;  DB 2;  Length 1291;
  Best Local Similarity   23.6%;  Pred. No. 1.3;
  Matches   37;  Conservative   21;  Mismatches   77;  Indels   22;  Gaps    4;

Qy       2 NSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRL--------ATNASQAGV 53
           |:    | | |||:       |  ||  | :| :: : || | |         | | |   :
Db    1139 NNDFNEGYKIIIKRIRGNTNDTRVRGGDILYFDMTINNKAYNLFMKNETMYADNHSTEDI 1198

Qy      54 EKILSALEIPDVGNLSQVVVMKSKNDQGITNKC-KMNLQDNNGNDIGFIGFHQF------ 106
             |    :  |: :     :    |    ::  | |    | : |  || ::|
Db    1199 YAIGLREQTKDINDNIIFQIQPMNNTYYYASQIFKSNFNGENISGICSIGTYRFRLGGDW 1258

Qy     107 ---NNIAKLVASNWYNRQIERSSRTLGCSWEFIPVDD 140
             | :   |    |   :| :|      | |:|| :
Db    1259 YRHNYLVPTVKQGNYASLLESTS----THWGFVPVSE 1291


Search completed: March 13, 2003, 11:40:13
Job time : 23.4798 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:38:07 ; Search time 10.8934 Seconds
(without alignments)
1235.505 Million cell updates/sec

Title: US-09-917-791-22
Perfect score: 727
Sequence: 1 LNSSLYRGTKFIIKKYASGN.....QIERSSRTLGCSWEFIPVDD 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	727	100.0	1296	1	BTCLAB	bontoxilysin (EC 3
2	671	92.3	1296	2	I40645	botulinum neurotox
3	282.5	38.9	1268	2	S33411	botulinum neurotox
4	258.5	35.6	1274	2	I40813	neurotoxin type F
5	206.5	28.4	1252	2	S21178	botulinum neurotox
6	190	26.1	1251	2	JH0256	botulinum neurotox
7	169.5	23.3	1291	2	I40631	non-proteolytic bo
8	165.5	22.8	1291	1	A48940	bontoxilysin (EC 3
9	163	22.4	1315	1	BTCLTN	tentoxilysin (EC 3
10	158.5	21.8	1297	2	S39791	neurotoxin - Clost
11	110	15.1	315	2	C97224	hypothetical prote
12	100	13.8	1291	2	S46431	botulinum neurotox
13	100	13.8	1291	2	A49777	botulinum neurotox
14	95	13.1	1285	2	S70582	botulinum neurotox
15	85.5	11.8	467	2	D96901	probable dehydroge
16	83	11.4	4550	2	T18440	hypothetical prote
17	79.5	10.9	233	2	AI1580	ABC transporter, A
18	79	10.9	122	2	A53878	type E neurotoxin
19	79	10.9	501	2	T43047	retrovirus-related
20	79	10.9	3394	2	T18501	hypothetical prote
21	78.5	10.8	319	2	D97066	periplasmic sugar-
22	78.5	10.8	528	2	T50330	hypothetical prote
23	78.5	10.8	659	2	AB1293	ABC transporter (p
24	78.5	10.8	768	2	T18461	hypothetical prote
25	78	10.7	145	2	AB2142	type I restriction
26	78	10.7	385	2	C97277	glycosyltransferas
27	77.5	10.7	236	2	H98058	hypothetical prote
28	77.5	10.7	504	2	D71615	hypothetical prote
29	77	10.6	271	2	D82314	transcription regu
30	77	10.6	4981	2	T18489	hypothetical prote
31	76.5	10.5	993	2	B64695	type I restriction
32	76.5	10.5	1278	2	B70236	hypothetical prote
33	76	10.5	472	2	AD2284	hypothetical prote
34	76	10.5	665	2	B71609	hypothetical prote
35	76	10.5	1175	2	F64489	hypothetical prote
36	75.5	10.4	233	2	AG1227	ABC transporter, A
37	75.5	10.4	323	2	C83940	sugar ABC transpor
38	75	10.3	257	2	A96908	ABC transporter, A
39	75	10.3	708	2	A95214	cell wall surface
40	75	10.3	719	2	B98078	hypothetical prote
41	75	10.3	732	2	C64972	probable ATPase -
42	75	10.3	1247	2	E71616	hypothetical prote
43	75	10.3	1711	2	T18429	hypothetical prote
44	74.5	10.2	233	2	C95192	ABC transporter, A
45	74.5	10.2	435	2	AI1017	probable type-I se

ALIGNMENTS

RESULT 1
BTCLAB
bontoxilysin (EC 3.4.24.69) - A precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type A
C;Species: Clostridium botulinum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: A35294
A;Molecule type: DNA
A;Residues: 1-1296 <BIN>
A;Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
A;Experimental source: strain 62A, subtype A
R;Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990
A;Title: The complete amino acid sequence of the Clostridium botulinum type A neuroto
A;Reference number: S09492; MUID:90235864; PMID:2185020
A;Accession: S09492
A;Molecule type: DNA
A;Residues: 1,'Q',3-26,'V',28-1296 <THO>
A;Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382
A;Experimental source: NCTC 2916
R;Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone
A;Reference number: S67988; MUID:96096783; PMID:8521962
A;Accession: S68220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <FUJ>
A;Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
R;Betley, M.J.; Somers, E.; DasGupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A;Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t
A;Reference number: A33401; MUID:89350959; PMID:2669749
A;Accession: A33401
A;Molecule type: DNA
A;Residues: 1-35 <BET>
A;Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
R;Gimenez, J.A.; DasGupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A;Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
A;Reference number: A53884; MUID:94000342; PMID:8397793
A;Accession: A53884
A;Status: preliminary
A;Molecule type: protein
A;Residues: 867-880;1148-1217,'Y',1219 <GIM>
A;Experimental source: strain Hall
A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after

```
R;DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and arou
A;Reference number: A60025; MUID:91120847; PMID:2126206
A;Accession: A60025
A;Molecule type: protein
A;Residues: 2-6;445-453,'X',455-457 <DAS1>
R;DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A;Reference number: A27000
A;Accession: A27000
A;Molecule type: protein
A;Residues: 2-47 <DAS2>
R;Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A;Title: Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.
A;Reference number: A49708; MUID:94124495; PMID:8294407
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
C;Genetics:
A;Gene: atx; botA
C;Function:
A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
F;2-444/Product: bontoxilysin A light chain #status experimental <LGHT>
F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
F;223,227/Binding site: zinc (His) #status predicted
F;224/Active site: Glu #status predicted

  Query Match             100.0%;  Score 727;  DB 1;  Length 1296;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-58;
  Matches  140;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1150 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 1209

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1210 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 1269

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           ||||||||||||||||||||
Db    1270 QIERSSRTLGCSWEFIPVDD 1289


RESULT 2
I40645
botulinum neurotoxin type A - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40645
R;Willems, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A;Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A;Reference number: I40645; MUID:94143603; PMID:8310180
A;Accession: I40645
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1296 <RES>
A;Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             92.3%;  Score 671;  DB 2;  Length 1296;
  Best Local Similarity   90.7%;  Pred. No. 7.3e-53;
  Matches  127;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 60
           |||:|| |||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db    1150 LNSTLYEGTKFIIKKYASGNEDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSAL 1209

Qy      61 EIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWYNR 120
           ||||||||||||||||||:|||| |||||||||||||||||||| ::|||||||||||||
Db    1210 EIPDVGNLSQVVVMKSKDDQGIRNKCKMNLQDNNGNDIGFIGFHLYDNIAKLVASNWYNR 1269

Qy     121 QIERSSRTLGCSWEFIPVDD 140
           |: ::|||| ||||||||||
Db    1270 QVGKASRTFGCSWEFIPVDD 1289


RESULT 3
S33411
botulinum neurotoxin type F - Clostridium barati
C;Species: Clostridium barati
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S33411; S31860
R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P
FEMS Microbiol. Lett. 108, 175-182, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neuroto
A;Reference number: S33411; MUID:93252228; PMID:8486245
A;Accession: S33411
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1268 <THO>
A;Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             38.9%;  Score 282.5;  DB 2;  Length 1268;
  Best Local Similarity   45.3%;  Pred. No. 1.5e-17;
  Matches   63;  Conservative   19;  Mismatches   44;  Indels   13;  Gaps    5;

Qy       2 NSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILS 58
           |: || | : ||:|  |    | || || || ||||||  | ||:|  : | : ||| :
Db    1131 NARLYTGVEVIIRKVGSTDTSNTDNFVRKNDTVYINVVDGNSEYQLYADVSTSAVEKTIK 1190

Qy      59 ALEIPDVG-NLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
             | :   | :|:::| |     | : | || : ||||||| :|||  ||   ||||:|
Db    1191 LRRISNSNYNSNQMIIMDS-----IGDNCTMNFKTNNGNDIGLLGFH-LNN---LVASSW 1241

Qy     118 YNRQIERSSRTLGCSWEFI 136
           | : |  ::|  || | ||
Db    1242 YYKNIRNNTRNNGCFWSFI 1260


RESULT 4
I40813
neurotoxin type F - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40813; S48108
R;East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson,
FEMS Microbiol. Lett. 96, 225-230, 1992
A;Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A;Reference number: I40644
A;Accession: I40813
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1274 <RES>
A;Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48108
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 634-1002 <CAM>
A;Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             35.6%;  Score 258.5;  DB 2;  Length 1274;
```

```
  Best Local Similarity   43.2%;  Pred. No. 2.4e-15;
  Matches   60;  Conservative   20;  Mismatches   44;  Indels   15;  Gaps    5;

Qy       1 LNSSLYRGTKFIIKKYAS---GNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKIL 57
           ||  || | : ||:|       | || || ||  ||||| :  ||||  : :::  |||:
Db    1140 LNYKLYEGVEVIIRKNGPIDISNTDNFVRKNDLAYINVVDRGVEYRLYAD-TKSEKEKII 1198

Qy      58 SALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
             : |  :| |::|| |     | | | || |:|||::|| :|||  |    ||||:|
Db    1199 RTSNLND--SLGQIIVMDS-----IGNNCTMNFQNNNGSNIGLLGFHSNN----LVASSW 1247

Qy     118 YNRQIERSSRTLGCSWEFI 136
           |   | |:: : || |  |
Db    1248 YYNNIRRNTSSNGCFWSSI 1266


RESULT 5
S21178
botulinum neurotoxin type E precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
C;Accession: S21178; S48107; JH0257; B35294; A60027; S18111
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992
A;Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
A;Reference number: S21178; MUID:92174922; PMID:1541280
A;Accession: S21178
A;Molecule type: DNA
A;Residues: 1-1252 <WHE>
A;Cross-references: EMBL:X62683; NID:g40397; PIDN:CAA44558.1; PID:g40398
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48107
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 616-982 <CAM>
A;Cross-references: EMBL:X70815; NID:g407786; PIDN:CAA50146.1; PID:g407787
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum type
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0257
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-176,'R',178-197,'C',199-339,'R',341-772,'I',774-962,'FE',965-966,'R',968-1
A;Cross-references: EMBL:X62089; NID:g40393; PIDN:CAA43999.1; PID:g40394
A;Experimental source: strain Beluga
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: B35294
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-176,'R',178-252 <BIN>
A;Experimental source: strain Beluga
R;Gimenez, J.A.; DasGupta, B.R.
Biochimie 72, 213-217, 1990
A;Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the si
A;Reference number: A60027; MUID:90344918; PMID:2116911
A;Accession: A60027
A;Molecule type: protein
A;Residues: 420-427 <GIM>
A;Experimental source: strain Beluga
A;Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F;423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
F;412-426/Disulfide bonds: #status predicted

  Query Match           28.4%;  Score 206.5;  DB 2;  Length 1252;
  Best Local Similarity   38.1%;  Pred. No. 1.2e-10;
  Matches   53;  Conservative   20;  Mismatches   51;  Indels   15;  Gaps    6;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVV-KNKEYRLATNASQAGVEKILS 58
             | : || | |  |::   |   ||:|| ||:|||| |  |   : |  : :    ||
Db    1117 LANRLYSGIKVKIQRVNNSSTNDNLVRKNDQVYINFVASKTHLFPLYADTATTNKEK--- 1173

Qy      59 ALEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNW 117
             ::|   ||  :|||| |     : | | || ::||||:|| :||        :||| |
Db    1174 TIKISSSGNRFNQVVVMNS-----VGNNCTMNFKNNNGNNIGLLGF----KADTVVASTW 1224

Qy     118 YNRQIERSSRTLGCSWEFI 136
           |   :   : : || | ||
Db    1225 YYTHMRDHTNSNGCFWNFI 1243


RESULT 6
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C;Accession: JH0256; S16145
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum t
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0256
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-27,'E',29-1251 <POU>
A;Cross-references: EMBL:X62088; NID:g40379
A;Experimental source: strains ATCC 43181 and ATCC 43755
R;Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa,
J. Gen. Microbiol. 137, 519-525, 1991
A;Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E t
A;Reference number: S16145; MUID:91237316; PMID:2033376
A;Accession: S16145
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229,'M',231-252 <FUJ>
A;Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408
A;Experimental source: strain BL6340
C;Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter relea
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F;423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F;412-426/Disulfide bonds: #status predicted

  Query Match           26.1%;  Score 190;  DB 2;  Length 1251;
  Best Local Similarity   37.7%;  Pred. No. 3.9e-09;
  Matches   52;  Conservative   19;  Mismatches   53;  Indels   14;  Gaps    6;

Qy       1 LNSSLYRGTKFIIKKY-ASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSA 59
             | : || | |  |::   |   ||:|| ||:|||| |  :| : |   |  |   |
Db    1117 LANRLYSGIKVKIQRVNNSSTNDNLVRKNDQVYINFVA-SKTHLLPLYADTATTNK-EKT 1174

Qy      60 LEIPDVGN-LSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNIAKLVASNWY 118
           ::|   ||  :|||| |         | || ::||||:|| :||        :||| ||
Db    1175 IKISSSGNRFNQVVVMNS------VGNCTMNFKNNNGNNIGLLGF----KADTVVASTWY 1224

Qy     119 NRQIERSSRTLGCSWEFI 136
             :  :: : |  | ||
Db    1225 YTHMRDNTNSNGFFWNFI 1242


RESULT 7
I40631
```

non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40631; S48103; S48104; S36015
R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinu
A;Reference number: I40631; MUID:94122659; PMID:7764370
A;Accession: I40631
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1291 <RES>
A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761,'E',763-841,'M',843,'T',845,'N',847-994 <CAM1>
A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A;Experimental source: non-proteolytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843,'T',845,'N',847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C;Comment: Botulinum neurotoxin type B in these strains may posses a capable catalytic s
C;Genetics:
A;Gene: bont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

Query Match 23.3%; Score 169.5; DB 2; Length 1291;
Best Local Similarity 28.3%; Pred. No. 3e-07;
Matches 43; Conservative 39; Mismatches 53; Indels 17; Gaps 7;

Qy 4 SLYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRL-ATNASQAGVEKILSAL 60
Db 1138 NLYIGEKFIIRRESNSQSINDDIVRKEDYIHLDLVLHHEEWRVYAYKYFKEQEEKLFLSI 1197

Qy 61 EIPDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIAKL------ 112
Db 1198 -ISDSNEFYKTIEIKEYDEQP-SYSCQLLFKKDEESTDDIGLIGIHRFYESGVLRKKYKD 1255

Qy 113 --VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
Db 1256 YFCISKWYLKEVKRKPYKSNLGCNWQFIPKDE 1287

RESULT 8
A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type B (BoNT/B)
C;Species: Clostridium botulinum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the typ
A;Reference number: A48940; MUID:92384550; PMID:1514783
A;Accession: A48940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <WHE>
A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A;Experimental source: type B, Danish
A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publ
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-994 <CAM>
A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A;Experimental source: proteolytic type B, strain NCTC 7273
R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A;Description: Partial amino acid sequence of botulinum neurotoxin type B and compari
A;Reference number: S21575
A;Accession: S21575
A;Molecule type: DNA
A;Residues: 36-217,'G',219-224,'S',226-246 <SZA>
A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars
J. Biol. Chem. 267, 14721-14729, 1992
A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus
A;Reference number: A42871; MUID:92340509; PMID:1634516
A;Accession: A42871
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-313,'S',315-451 <KUR>
A;Experimental source: strain Okra
A;Note: sequence extracted from NCBI backbone (NCBIP:109365)
R;DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity wi
A;Reference number: S07155; MUID:89000987; PMID:3139097
A;Accession: S07155
A;Molecule type: protein
A;Residues: 2-29,'M',31-45 <DAS>
A;Accession: S08562
A;Molecule type: protein
A;Residues: 442-463,'R',465-467 <DA2>
R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A;Reference number: S07128; MUID:85197963; PMID:3888113
A;Accession: S07128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-16 <SCH1>
A;Accession: S08573
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-17 <SCH2>
A;Accession: S08574
A;Status: preliminary
A;Molecule type: protein
A;Residues: 442-459 <SCH3>
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C;Genetics:
A;Gene: bont/b
C;Function:
A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

```
  Query Match             22.8%;  Score 165.5;  DB 1;  Length 1291;
  Best Local Similarity   28.7%;  Pred. No. 6.9e-07;
  Matches   43;  Conservative   31;  Mismatches   61;  Indels   15;  Gaps    5;

Qy       5 LYRGTKFIIKKYASGN--KDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEI 62
           || | ||||:: ::     |:|||  | :|::    |:|:|: |       |: |    |
Db    1139 LYIGEKFIIRRKSNSQSINDDIVRKEDYIYLDFFNLNQEWRVYTYKYFKKEEEKLFLAPI 1198

Qy      63 PDVGNLSQVVVMKSKNDQGITNKCKM--NLQDNNGNDIGFIGFHQFNNIA--------KL 112
             |      : :|  ::|  |  |::     : : ::|| || |:|
Db    1199 SDSDEFYNTIQIKEYDEQP-TYSCQLLFKKDEESTDEIGLIGIHRFYESGIVFEEYKDYF 1257

Qy     113 VASNWYNRQIERS--SRTLGCSWEFIPVDD 140
           | || ::::|   :  |||:|:||| |:
Db    1258 CISKWYLKEVKRKPYNLKLGCNWQFIPKDE 1287


RESULT 9
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botu
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A;Experimental source: strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation; epitope region
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.R
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolyt
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A;Reference number: S69348; MUID:95262688; PMID:7744050
A;Accession: S69348
A;Molecule type: protein
A;Residues: 2-31 <DEF>
C;Comment: The source of this protein was an extrachromosomal plasmid.
C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (
dual chains are not toxic when separated). The amino end of the heavy chain (fragment
C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to gang
C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internali
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of
C;Function:
A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in syn
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-457/Product: tentoxylysin light chain (fragment A) #status predicted <TTL>
F;461-1315/Product: tentoxylysin heavy chain (fragment B.C) #status experimental <TTH
F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F;233,237/Binding site: zinc (His) #status predicted
F;234/Active site: Glu #status predicted

  Query Match             22.4%;  Score 163;  DB 1;  Length 1315;
  Best Local Similarity   26.9%;  Pred. No. 1.2e-06;
  Matches   43;  Conservative   26;  Mismatches   49;  Indels   42;  Gaps    6;

Qy       5 LYRGTKFIIKKYASGNK-DNIVRNNDRVYINVVVKNKEY-----------------RLA 45
           || | |||||:|   |: |: |:: | : : |   | |:                  |:
Db    1169 LYNGLKFIIKRYTPNNEIDSFVKSGDFIKLYVSYNNNEHIVGYPKDGNAFNNLDRILRVG 1228

Qy      46 TNASQAGVEKILSALEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFH- 104
            ||    : | : |::: |:                  |   :: | |:    :| :| |
Db    1229 YNAPGIPLYKKMEAVKLRDLK---------------TYSVQLKLYDDKNASLGLVGTHN 1272

Qy     105 -QFNNIAK---LVASNWYNRQIERSSRTLGCSWEFIPVDD 140
            |  |      |:|||||   ::   : ||| | |:| |:
Db    1273 GQIGNDPNRDILIASNWYFNHLK--DKILGCDWYFVPTDE 1310


RESULT 10
S39791
neurotoxin - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C;Accession: S39791
R;Campbell, K.; Collins, M.D.; East, A.K.
Biochim. Biophys. Acta 1216, 487-491, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridiu
A;Reference number: S39791; MUID:94092745; PMID:8268233
A;Accession: S39791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1297 <CAM>
A;Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             21.8%;  Score 158.5;  DB 2;  Length 1297;
  Best Local Similarity   29.9%;  Pred. No. 3e-06;
  Matches   46;  Conservative   28;  Mismatches   59;  Indels   21;  Gaps    7;

Qy       4 SLYRGTKFIIKKYASG---NKDNIVRNNDRVYINV-VVKNKEYRLATNASQAGVEKILSA 59
           :|| | :|||||  ::    | |||||  | :|:|:  : :: ||:    :   ::  |
Db    1144 NLYLGLRFIIKKASNSRNINNDNIVREGDYIYLNIDNISDESYRVYVLVNSKEIQTQLFL 1203

Qy      60 LEIPDVGNLSQVVVMKSKNDQGITNKCKMNLQDNNGNDIGFIGFHQF----------NN 108
             | |      |: :| |  :  |  |:: | : :    |  |  :|          :|
Db    1204 APINDDPTFYDVLQIK-KYYEKTTYNCQI-LCEKDTKTFGLFGIGKFVKDYGYVWDTYDN 1261

Qy     109 IAKLVASNWYNRQIER--SSRTLGCSWEFIPVDD 140
                 | || |:|    :   |||:|:|||||:
Db    1262 Y--FCISQWYLRRISENINKLRLGCNWQFIPVDE 1293


RESULT 11
```

```
C97224
hypothetical protein CAC2635 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: C97224
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97224
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-315 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80582.1; PID:g15025662; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2635

  Query Match             15.1%;  Score 110;  DB 2;  Length 315;
  Best Local Similarity   32.6%;  Pred. No. 0.014;
  Matches   44;  Conservative   21;  Mismatches   52;  Indels   18;  Gaps    8;

Qy      12 IIKKY---ASGNKDNIV--RNNDRVYINVVVKNKEYRL--ATNASQAGVEKILSALE--I 62
           |||||     | ||||    ||::  | |  |:| | |  :   :  |  |  | |:  :
Db      47 IIKKYDVTGDGLKDNIYIKINNNKYSIAVKSKDKIYTLQPSHKLNSLGTYKSYSPLKLIL 106

Qy      63 PDVG--NLSQVVVMKSKNDQG-----ITNKCKM-NLQDNNGNDIGFIGFHQFNNIAKLVA 114
             ||   |:|::   |:|:       | :| |  ::    | ||||  |  |   |::: 
Db     107 TDVSRDNISELFTQSSENNTNLQHLFIWDKTKFKDILHTTNNIIGFIDVHN-NKTPKIIS 165

Qy     115 SNWYNRQIERSSRTL 129
           ||: |  :  |:  |
Db     166 SNFNNNDLVFSNYIL 180


RESULT 12
S46431
botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)
N;Alternate names: BoNT/C1 protein
C;Species: Clostridium botulinum phage 1C
A;Variety: strain C 468
C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
C;Accession: S46431; S49107
R;Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.
Mol. Gen. Genet. 243, 631-640, 1994
A;Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxic p
A;Reference number: S46426; MUID:94301293; PMID:8028579
A;Accession: S46431
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1291 <HAU>
A;Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175
A;Experimental source: strain C 468
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C;Superfamily: tetanus toxin

  Query Match             13.8%;  Score 100;  DB 2;  Length 1291;
  Best Local Similarity   23.6%;  Pred. No. 0.62;
  Matches   37;  Conservative   21;  Mismatches   77;  Indels   22;  Gaps    4;

Qy       2 NSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRL--------ATNASQAGV 53
           |:    | | |||:      |  ||  | :| :: : || | |        | | |   :
Db    1139 NNDFNEGYKIIIKRIRGNTNDTRVRGGDILYFDMTINNKAYNLFMKNETMYADNHSTEDI 1198

Qy      54 EKILSALEIPDVGNLSQVVVMKSKNDQGITNKC-KMNLQDNNGNDIGFIGFHQF------ 106
             |    :  |: :      :    |     ::  | :  |  | : |  || ::|
Db    1199 YAIGLREQTKDINDNIIFQIQPMNNTYYYASQIFKSNFNGENISGICSIGTYRFRLGGDW 1258

Qy     107 ---NNIAKLVASNWYNRQIERSSRTLGCSWEFIPVDD 140
             | :   |    |   :| :|      | |:|| :
Db    1259 YRHNYLVPTVKQGNYASLLESTS----THWGFVPVSE 1291


RESULT 13
A49777
botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain
C;Species: Clostridium botulinum phage
C;Date: 10-Mar-1994 #sequence_revision 07-Apr-1994 #text_change 23-Mar-2001
C;Accession: S11291; A35396; S22166; A49777
R;Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet,
Nucleic Acids Res. 18, 4924, 1990
A;Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
A;Reference number: S11291; MUID:90370487; PMID:2204031
A;Accession: S11291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84,'P',86-1291 <HAU>
A;Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi,
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A;Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 t
A;Reference number: A35396; MUID:91024998; PMID:2222445
A;Accession: A35396
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-669,'R',671-1291 <TS1>
R;Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
submitted to the EMBL Data Library, December 1991
A;Description: Nucleotide sequence of the gene for one of the components of hemagglut
A;Reference number: S22163
A;Accession: S22166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <TS2>
A;Cross-references: EMBL:X62389; NID:g558175; PIDN:CAA44263.1; PID:g40390
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K
Appl. Environ. Microbiol. 57, 1168-1172, 1991
A;Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and
A;Reference number: A49777; MUID:91282468; PMID:2059039
A;Accession: A49777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-607 <TS3>
A;Cross-references: GB:D90210
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             13.8%;  Score 100;  DB 2;  Length 1291;
  Best Local Similarity   23.6%;  Pred. No. 0.62;
  Matches   37;  Conservative   21;  Mismatches   77;  Indels   22;  Gaps    4;

Qy       2 NSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRL--------ATNASQAGV 53
           |:    | | |||:      |  ||  | :| :: : || | |        | | |   :
Db    1139 NNDFNEGYKIIIKRIRGNTNDTRVRGGDILYFDMTINNKAYNLFMKNETMYADNHSTEDI 1198

Qy      54 EKILSALEIPDVGNLSQVVVMKSKNDQGITNKC-KMNLQDNNGNDIGFIGFHQF------ 106
             |    :  |: :      :    |     ::  | :  |  | : |  || ::|
Db    1199 YAIGLREQTKDINDNIIFQIQPMNNTYYYASQIFKSNFNGENISGICSIGTYRFRLGGDW 1258

Qy     107 ---NNIAKLVASNWYNRQIERSSRTLGCSWEFIPVDD 140
             | :   |    |   :| :|      | |:|| :
Db    1259 YRHNYLVPTVKQGNYASLLESTS----THWGFVPVSE 1291


RESULT 14
S70582
botulinum neurotoxin type Dsa precursor - Clostridium botulinum phage d-sA
C;Species: Clostridium botulinum phage d-sA
A;Note: host Clostridium botulinum type D (strain South Africa)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C;Accession: S70582
R;Morilishi, K.; Koura, M.; Abe, N.; Fujii, N.; Fujinaga, Y.; Inoue, K.; Ogumad, K.
Biochim. Biophys. Acta 1307, 123-126, 1996
A;Title: Mosaic structures of neurotoxins produced from Clostridium botulinum types C
A;Reference number: S70582; MUID:96283801; PMID:8679691
```

```
A;Accession: S70582
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1285 <MOR>
A;Cross-references: EMBL:D38442; NID:g1374775; PIDN:BAA07477.1; PID:g1374776
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
a disulfide bond. The heavy chain mediates the binding of toxin to the presynaptic membr
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; neurotoxin; transmembrane protein
F;1-447/Product: botulinum neurotoxin type Dsa light chain #status predicted <MAT1>
F;448-1285/Product: botulinum neurotoxin type Dsa heavy chain #status predicted <MAT2>

  Query Match             13.1%;  Score 95;  DB 2;  Length 1285;
  Best Local Similarity   23.8%;  Pred. No. 1.8;
  Matches   38;  Conservative   20;  Mismatches   72;  Indels   30;  Gaps    7;

Qy       2 NSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALE 61
           |:    | | |||:      |  ||  : :| |  : ||:| |       :       :
Db    1135 NNDFNEGYKIIIKRIRGNTNDTRVRGENVLYFNTTIDNKQYSLGMYKPSRNL-----GTD 1189

Qy      62 IPDVGNLSQVVVMKSKNDQGITNKC--------KMNLQDN-NGNDIGF--IGFHQFNNIA 110
           :  :| | | :    |    |   |         :: |  |   | :|   || : |
Db    1190 LVPLGALDQPMDEIRKYGSFIIQPCNTFDYYASQLFLSSNATTNRLGILSIGSYSF---- 1245

Qy     111 KLVASNWYNRQ-------IERSSRTL---GCSWEFIPVDD 140
           ||    |:| :       ||  :  |      | |:|  :
Db    1246 KLGDDYWFNHEYLIPVIKIEHYASLLESTSTHWVFVPASE 1285


RESULT 15
D96901
probable dehydrogenase with iron-sulfur domain [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: D96901
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D96901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK77999.1; PID:g15022830; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0012

  Query Match             11.8%;  Score 85.5;  DB 2;  Length 467;
  Best Local Similarity   28.8%;  Pred. No. 3.9;
  Matches   40;  Conservative   25;  Mismatches   51;  Indels   23;  Gaps    8;

Qy       1 LNSSLYRGTKFIIKKYASGNKDNIVRNNDRVYINVVVK-NKEYRLATNASQAGVEKILSA 59
           || ||      | |:|  |:  :|| |:| ||  :  | |  | | |  :|  :||| |
Db      26 LNVSL------IEKEYEIGDDTSIV-NSDLVYDGIESKYNFNYGLETEGNQE-IEKICSV 77

Qy      60 LEIPDVGNLSQVVVMKSKNDQGITNK-----CKMNLQDNNGNDIGFIGFHQFNNIAK-LV 113
           | :|    :: :|  : | ::   ||      |  ::| :  :|        : ||| ::
Db      78 LSVP-FKKVNAIVCSRDKENEDTINKIYSRGIKRGIKDASLANIE-------SAIAKSII 129

Qy     114 ASNWYNRQIERSSRTLGCS 132
           :  |: ||     : |:
Db     130 LKDINNKGIELKDIGIVCA 148


Search completed: March 13, 2003, 11:41:46
Job time : 12.8934 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:36:37 ; Search time 9.54467 Seconds
(without alignments)
899.518 Million cell updates/sec

Title: US-09-917-791-21
Perfect score: 1071
Sequence: 1 IKVNNWDLFFSPSEDNFTND.....NIGNMLYKDDFVGALIFSGA 207

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1071	100.0	1295	1	BXA1_CLOBO	P10845 clostridium
2	918	85.7	1295	1	BXA2_CLOBO	Q45894 clostridium
3	377	35.2	1296	1	BXG_CLOBO	Q60393 clostridium
4	374.5	35.0	1250	1	BXE_CLOBU	P30995 clostridium
5	368.5	34.4	1250	1	BXE_CLOBO	Q00496 clostridium
6	366.5	34.2	1290	1	BXB_CLOBO	P10844 clostridium
7	316	29.5	1274	1	BXF_CLOBO	P30996 clostridium
8	312	29.1	1314	1	TETX_CLOTE	P04958 clostridium
9	290.5	27.1	1276	1	BXD_CLOBO	P19321 clostridium
10	263	24.6	1290	1	BXC1_CLOBO	P18640 clostridium
11	107	10.0	246	1	Y402_BUCAI	P57482 buchnera ap
12	106	9.9	1196	1	BXCN_CLOBO	P46081 clostridium
13	98	9.2	345	1	VA0D_YEAST	P32366 saccharomyc
14	95.5	8.9	1032	1	MT18_YEAST	P40469 saccharomyc
15	93.5	8.7	355	1	Y198_RICPR	Q9zdw7 rickettsia
16	91	8.5	635	1	HTPG_VIBCH	P22359 vibrio chol
17	91	8.5	1162	1	BXEN_CLOBU	Q06366 clostridium
18	90.5	8.5	1186	1	CAGA_HELPY	P55980 helicobacte
19	89	8.3	1162	1	BXEN_CLOBO	P46082 clostridium
20	88.5	8.3	556	1	SYR_LISIN	Q927t2 listeria in
21	87	8.1	755	1	KHL5_HUMAN	Q96pq7 homo sapien
22	86.5	8.1	556	1	SYR_LISMO	Q8y493 listeria mo
23	86.5	8.1	1939	1	MYH4_HUMAN	Q9y623 homo sapien
24	86	8.0	610	1	MUTL_BORBU	O51229 borrelia bu
25	86	8.0	844	1	SECA_STACA	P47994 staphylococ
26	86	8.0	891	1	SECA_PAVLU	Q01570 pavlova lut
27	86	8.0	974	1	TRP4_MOUSE	Q9quq5 mus musculu
28	86	8.0	977	1	TRP4_HUMAN	Q9ubn4 homo sapien
29	86	8.0	977	1	TRP4_RAT	O35119 rattus norv
30	86	8.0	981	1	TRP4_BOVIN	P79100 bos taurus
31	85.5	8.0	426	1	HEM1_SULSO	Q980u7 sulfolobus
32	85.5	8.0	1167	1	CAGA_HELPJ	Q9zlt1 helicobacte
33	85	7.9	336	1	Y05G_BPT4	P39242 bacteriopha
34	85	7.9	1273	1	YAR2_SCHPO	Q10135 schizosacch
35	85	7.9	1938	1	MYSS_CHICK	P13538 gallus gall
36	84.5	7.9	2493	1	YBA4_YEAST	P35194 saccharomyc
37	83	7.7	190	1	PBP_PLAFA	P54189 plasmodium
38	82.5	7.7	617	1	Y237_BUCAI	P57331 buchnera ap
39	82.5	7.7	872	1	TE95_TETTH	Q94819 tetrahymena
40	82.5	7.7	1058	1	CARB_FUSNN	Q8rg86 fusobacteri
41	82.5	7.7	1228	1	ATC_PLAFK	Q08853 plasmodium
42	82.5	7.7	1941	1	MYH2_HUMAN	Q9ukx2 homo sapien
43	82.5	7.7	1947	1	MYSC_CAEEL	P12845 caenorhabdi
44	82	7.7	505	1	Y038_BORBU	O51067 borrelia bu
45	82	7.7	528	1	YE19_SCHPO	O13874 schizosacch

ALIGNMENTS

RESULT 1
BXA1_CLOBO
ID BXA1_CLOBO STANDARD; PRT; 1295 AA.
AC P10845; P18639; P01561;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 2916;
RX MEDLINE=90235864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
RA Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT gene.";
RL Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=62A;
RX MEDLINE=90264400; PubMed=2160960;
RA Binz B., Kuarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=62A;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=Hall;
RX MEDLINE=89350959; PubMed=2669749;
RA Betley M.J., Somers E., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
RT the N-terminal encoding region.";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=Type A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;

```
RT   "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT   components of Clostridium botulinum type A progenitor toxins.";
RL   FEBS Lett. 376:41-44(1995).
RN   [6]
RP   SEQUENCE OF 1-16.
RX   MEDLINE=84178501; PubMed=6370252;
RA   Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequence of the heavy and light chains of
RT   botulinum neurotoxin type A.";
RL   Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN   [7]
RP   SEQUENCE OF 1-46.
RA   Dasgupta B.R., Foley J., Niece R.;
RT   "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL   Biochemistry 26:4162-4162(1987).
RN   [8]
RP   SEQUENCE OF 1-5 AND 444-456.
RX   MEDLINE=91120847; PubMed=2126206;
RA   Dasgupta B.R., Dekleva M.L.;
RT   "Botulinum neurotoxin type A: sequence of amino acids at the
RT   N-terminus and around the nicking site.";
RL   Biochimie 72:661-664(1990).
RN   [9]
RP   SEQUENCE OF 448-464 AND 872-895.
RX   MEDLINE=89024662; PubMed=3178218;
RA   Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT   "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RT   halves and their partial sequences.";
RL   Arch. Biochem. Biophys. 266:142-151(1988).
RN   [10]
RP   SEQUENCE OF 448-482.
RX   MEDLINE=85285016; PubMed=3896784;
RA   Shone C.C., Hambleton P., Melling J.;
RT   "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RT   and purification of two tryptic fragments. Proteolytic action near
RT   the COOH-terminus of the heavy subunit destroys toxin-binding
RT   activity.";
RL   Eur. J. Biochem. 151:75-82(1985).
RN   [11]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [12]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
RN   [13]
RP   MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX   MEDLINE=21556941; PubMed=11700044;
RA   Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
RT   "Site-directed mutagenesis identifies active-site residues of the
RT   light chain of botulinum neurotoxin type a.";
RL   Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX   MEDLINE=98455071; PubMed=9783750;
RA   Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT   "Crystal structure of botulinum neurotoxin type A and implications
RT   for toxicity.";
RL   Nat. Struct. Biol. 5:898-902(1998).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
CC       the treatment of strabismus and blepharospasm associated with
CC       dystonia and cervical dystonia. Also used for the treatment of
CC       hemifacial spasm and a number of other neurological disorders
CC       characterized by abnormal muscle contraction.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   -!- DATABASE: NAME=BOTOX product information Web site;
CC       WWW="http://www.botox.com/index.jsp?hp&productinfo".
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 19 of February 2002;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt019.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52066; CAA36289.1; -.
DR   EMBL; M30196; AAA23262.1; -.
DR   EMBL; X92973; CAA63551.1; -.
DR   EMBL; D67030; BAA11051.1; -.
DR   EMBL; M27892; AAA23269.1; -.
DR   PIR; A35294; BTCLAB.
DR   PIR; S09492; S09492.
DR   PDB; 3BTA; 01-OCT-99.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   Pharmaceutical; 3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC).
FT   ACT_SITE    223    223
FT   METAL       226    226       ZINC (CATALYTIC).
FT   METAL       261    261       ZINC (CATALYTIC).
FT   DISULFID    429    453       INTERCHAIN.
FT   DISULFID   1234   1279
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
FT   VARIANT      26     26       V -> A.
FT   MUTAGEN     261    261       E->A: DRASTIC DECREASE IN ENZYMATIC
FT                                ACTIVITY.
FT   MUTAGEN     265    265       F->A: DECREASE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN     365    365       Y->A: DECREASE IN ENZYMATIC ACTIVITY.
FT   CONFLICT      1      1       P -> Q (IN REF. 1).
FT   CONFLICT    479    479       E -> P (IN REF. 9).
FT   CONFLICT    875    875       T -> L (IN REF. 8).
FT   CONFLICT    891    891       S -> K (IN REF. 8).
SQ   SEQUENCE   1295 AA;  149322 MW;  858342F754862579 CRC64;

  Query Match             100.0%;  Score 1071;  DB 1;  Length 1295;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.5e-78;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 513

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 573

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 633

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
           ||||||||||||||||||||||||||||
Db     634 PYIGPALNIGNMLYKDDFVGALIFSGA 660


RESULT 2
BXA2_CLOBO
ID   BXA2_CLOBO     STANDARD;      PRT;  1295 AA.
AC   Q45894; P77780;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=94143603; PubMed=8310180;
RA   Willems A., East A.K., Lawson P.A., Collins M.D.;
RT   "Sequence of the gene coding for the neurotoxin of Clostridium
RT   botulinum type A associated with infant botulism: comparison with
RT   other clostridial neurotoxins.";
RL   Res. Microbiol. 144:547-556(1993).
RN   [2]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure (By similarity).
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73423; CAA51824.1; -.
DR   EMBL; X87974; CAA61234.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    223    223       BY SIMILARITY.
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    453       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   1234   1279       BY SIMILARITY.
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
SQ   SEQUENCE   1295 AA;  149279 MW;  5DA04A13D98D6372 CRC64;

  Query Match           85.7%;  Score 918;  DB 1;  Length 1295;
  Best Local Similarity   85.4%;  Pred. No. 3.3e-66;
  Matches  176;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||||||||||||||||||||:| ||||:||||||||||||||||||||:|||||||
Db     454 IKVNNWDLFFSPSEDNFTNDLDKVEEITADTNIEAAEENISLDLIQQYYLTFDFDNEPEN 513

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNE 120
           |||||||||||||| ||||||||||||||||||||||||||||| ||| |||| |
Db     514 ISIENLSSDIIGQLEPMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGDSRIILTNSAEE 573

Qy     121 ALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIII 180
           ||| |:  |||||| ||||:||| || ||| | |:|||||||||:||:| |||||||||:
Db     574 ALLKPNVAYTFFSSKYVKKINKAVEAFMFLNWAEELVYDFTDETNEVTTMDKIADITIIV 633

Qy     181 PYIGPALNIGNMLYKDDFVGALIFSG 206
           |||||||||||| | :|| |:||:|
Db     634 PYIGPALNIGNMLSKGEFVEAIIFTG 659


RESULT 3
BXG_CLOBO
ID   BXG_CLOBO      STANDARD;      PRT;  1296 AA.
AC   Q60393;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
DE   (Bontoxilysin G).
GN   BOTG.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=113 / 30;
RX   MEDLINE=94092745; PubMed=8268233;
RA   Campbell K., Collins M.D., East A.K.;
RT   "Nucleotide sequence of the gene coding for Clostridium botulinum
```

```
RT   (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT   with other clostridial neurotoxins.";
RL   Biochim. Biophys. Acta 1216:487-491(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74162; CAA52275.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0     0       BY SIMILARITY.
FT   CHAIN         1     441     BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT   CHAIN       442    1296     BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT   METAL       229     229     ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230     230     BY SIMILARITY.
FT   METAL       233     233     ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    435     449     INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1296 AA;  149013 MW;  DC8E47E15F665C31 CRC64;

  Query Match             35.2%;  Score 377;  DB 1;  Length 1296;
  Best Local Similarity   41.3%;  Pred. No. 1.3e-22;
  Matches   88;  Conservative   34;  Mismatches   75;  Indels   16;  Gaps    3;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLD-LIQQYYLTFNFDNEPE 59
           | ||| ||||  ::|:|: || | | |  :|     | | |:| ||    |:   |   |
Db     450 IIVNNEDLFFIANKDSFSKDLAKAETIAYNTQNNTIENNFSIDQLILDNDLSSGIDLPNE 509

Qy      60 NISIENLSSDI-----IGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIAL 114
           |        ||     | |  |           ||  :|  ::| || || |      : |
Db     510 NTEPFTNFDDIDIPVYIKQSAL----------KKIFVDGDSLFEYLHAQTFPSNIENLQL 559

Qy     115 TNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIA 174
           |||:|:|| | ::||||||:: |:| |    |::|: ||: :: ||| |::: || ||::
Db     560 TNSLNDALRNNNKVYTFFSTNLVEKANTVVGASLFVNWVKGVIDDFTSESTQKSTIDKVS 619

Qy     175 DITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |::||||||||||:||   |::|  |    ||
Db     620 DVSIIIPYIGPALNVGNETAKENFKNAFEIGGA 652


RESULT 4
BXE_CLOBU
ID   BXE_CLOBU      STANDARD;      PRT;  1250 AA.
AC   P30995;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 43181, and ATCC 43755;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE OF 1-251 FROM N.A.
RC   STRAIN=BL6340;
RX   MEDLINE=91237316; PubMed=2033376;
RA   Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA   Yokosawa N., Yashiki T., Oguma K.;
RT   "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT   type E toxin gene from Clostridium butyricum strain BL6340.";
RL   J. Gen. Microbiol. 137:519-525(1991).
RN   [3]
RP   SEQUENCE OF 1-48.
RC   STRAIN=5262;
RA   Gimenez J., Foley J., Dasgupta B.R.;
RT   "Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT   partial sequence and comparison.";
RL   FASEB J. 2:A1750-A1750(1988).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62088; CAA43998.1; -.
DR   EMBL; X53180; CAA37321.1; -.
DR   PIR; JH0256; JH0256.
DR   PIR; S16145; S16145.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
```

```
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    229    229       K -> M (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143265 MW;  8171B5B2C2312857 CRC64;

  Query Match             35.0%;  Score 374.5;  DB 1;  Length 1250;
  Best Local Similarity   38.9%;  Pred. No. 1.9e-22;
  Matches   84;  Conservative   51;  Mismatches   60;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
           |::|| :|||  ||::: :| :|  :||    ||: | |        : : |  | || :
Db     426 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 477

Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           : |  :| | |:  |       :|  :   ||    :::::::  :| || ||:   |::
Db     478 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 533

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           : ||:|:: |||   ::||||||:::  |||  :|||:|:||:::|:: ||| | :: || |
Db     534 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVGWIQQVLVDFTTEANQKSTVD 593

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||:|::|||| ||||||   | :|  ||   ||
Db     594 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 629


RESULT 5
BXE_CLOBO
ID   BXE_CLOBO      STANDARD;      PRT;  1250 AA.
AC   Q00496;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beluga;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92174922; PubMed=1541280;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RT   "The complete amino acid sequence of the Clostridium botulinum type-E
RT   neurotoxin, derived by nucleotide-sequence analysis of the encoding
RT   gene.";
RL   Eur. J. Biochem. 204:657-667(1992).
RN   [3]
RP   SEQUENCE OF 1-251 FROM N.A.
RX   MEDLINE=90264400; PubMed=2160960;
RA   Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
RT   "The complete sequence of botulinum neurotoxin type A and comparison
RT   with other clostridial neurotoxins.";
RL   J. Biol. Chem. 265:9153-9158(1990).
RN   [4]
RP   SEQUENCE OF 1-13.
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [5]
RP   SEQUENCE OF 419-426.
RX   MEDLINE=90344918; PubMed=2116911;
RA   Gimenez J.A., Dasgupta B.R.;
RT   "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT   reveals the site trypsin nicks and homology with tetanus
RT   neurotoxin.";
RL   Biochimie 72:213-217(1990).
RN   [6]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC       181 BOND IN SNAP-25.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62089; CAA43999.1; -.
DR   EMBL; X62683; CAA44558.1; -.
DR   PIR; A60027; A60027.
DR   PIR; B35294; B35294.
DR   PIR; JH0257; JH0257.
DR   PIR; S08575; S08575.
DR   PIR; S18111; S18111.
DR   PIR; S21178; S21178.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
```

```
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    176    176       R -> G (IN REF. 2).
FT   CONFLICT    197    197       C -> S (IN REF. 2 AND 3).
FT   CONFLICT    339    339       R -> A (IN REF. 2).
FT   CONFLICT    772    772       I -> L (IN REF. 2).
FT   CONFLICT    962    963       FE -> LQ (IN REF. 2).
FT   CONFLICT    966    966       R -> A (IN REF. 2).
FT   CONFLICT   1194   1194       N -> NN (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143712 MW;  D9FCE26DDA041EB4 CRC64;

  Query Match             34.4%;  Score 368.5;  DB 1;  Length 1250;
  Best Local Similarity   38.4%;  Pred. No. 5.8e-22;
  Matches   83;  Conservative   51;  Mismatches   61;  Indels   21;  Gaps    7;

Qy       1 IKVNNWDLFFSPSEDNFTND-LNKGEEI----TSDTNIEAAEENISLDLIQQYYLTFNFD 55
           |::|| :|||  ||::: :| :|  :||    ||: | |        : : |  | || :
Db     426 IEINNGELFFVASENSYNDDNINTPKEIDDTVTSNNNYE--------NDLDQVILNFNSE 477

Qy      56 NEPENISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSR 111
           : |  :| | |:  |      :|  :   ||    :::::::  :| || ||:   |::
Db     478 SAP-GLSDEKLNLTIQND-AYIPKYD--SNGTSDIEQHDVNELNVFFYLDAQKVPEGENN 533

Qy     112 IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTD 171
           : ||:|:: |||   ::||||||:::  |||  :||:|: |::|:: ||| | :: || |
Db     534 VNLTSSIDTALLEQPKIYTFFSSEFINNVNKPVQAALFVSWIQQVLVDFTTEANQKSTVD 593

Qy     172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
           |||||:|::|||| ||||||   | :|  ||   ||
Db     594 KIADISIVVPYIGLALNIGNEAQKGNFKDALELLGA 629


RESULT 6
BXB_CLOBO
ID   BXB_CLOBO      STANDARD;      PRT;  1290 AA.
AC   P10844; P10843;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE   (Bontoxilysin B).
GN   BOTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92384550; PubMed=1514783;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA   Minton N.P.;
RT   "Molecular cloning of the Clostridium botulinum structural gene
RT   encoding the type B neurotoxin and determination of its entire
RT   nucleotide sequence.";
RL   Appl. Environ. Microbiol. 58:2345-2354(1992).
RN   [2]
RP   SEQUENCE OF 35-245 FROM N.A.
RC   STRAIN=NCTC 7273;
RA   Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 633-993 FROM N.A.
RC   STRAIN=NCTC 7273;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-44 AND 441-466.
RC   STRAIN=657;
RX   MEDLINE=89000987; PubMed=3139097;
RA   Dasgupta B.R., Datta A.;
RT   "Botulinum neurotoxin type B (strain 657): partial sequence and
RT   similarity with tetanus toxin.";
RL   Biochimie 70:811-817(1988).
RN   [5]
RP   SEQUENCE OF 1-16 AND 441-458.
RC   STRAIN=OKRA;
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93054694; PubMed=1429690;
RA   Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT   "Botulinum neurotoxins are zinc proteins.";
RL   J. Biol. Chem. 267:23479-23483(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC       SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81186; AAA23211.1; -.
DR   EMBL; Z11934; CAA77991.1; -.
DR   EMBL; X70817; CAA50148.1; -.
DR   PIR; S07128; S07128.
DR   PIR; S07155; S07155.
DR   PIR; S08562; S08562.
DR   PIR; S08573; S08573.
DR   PIR; S08574; S08574.
DR   PIR; A48940; A48940.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    440       BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
```

```
FT   CHAIN       441   1290       BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    445       INTERCHAIN (PROBABLE).
FT   CONFLICT     29     29       T -> M (IN REF. 4).
FT   CONFLICT    217    217       R -> G (IN REF. 2).
FT   CONFLICT    224    224       A -> S (IN REF. 2).
FT   CONFLICT    463    463       S -> R (IN REF. 4).
SQ   SEQUENCE   1290 AA;  150670 MW;  D21746E2C024DF43 CRC64;

  Query Match             34.2%;  Score 366.5;  DB 1;  Length 1290;
  Best Local Similarity   39.4%;  Pred. No. 8.8e-22;
  Matches   87;  Conservative   39;  Mismatches   62;  Indels   33;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
Db     446 IDVDNEDLFFIADKNSFSDDLSKNERIEYNT-------------QSNYIENDF---PIN 488

Qy      61 ISIENLSSDIIGQLEL--------------MPNIERFPNGKKYELDKYTMFHYLRAQEFE 106
Db     489 ELI--LDTDLISKIELPSENTESLTDFNVDVPVYEKQPAIKKIFTDENTIFQYLYSQTFP 546

Qy     107 HGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSE 166
Db     547 LDIRDISLTSSFDDALLFSNKVYSFFSMDYIKTANKVVEAGLFAGWVKQIVNDFVIEANK 606

Qy     167 VSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
Db     607 SNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFEIAGA 647


RESULT 7
BXF_CLOBO
ID   BXF_CLOBO      STANDARD;      PRT;  1274 AA.
AC   P30996;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE   (Bontoxilysin F).
GN   BOTF.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23387;
RX   MEDLINE=93012902; PubMed=1398040;
RA   East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA   Roberts T.A., Thompson D.E.;
RT   "Sequence of the gene encoding type F neurotoxin of Clostridium
RT   botulinum.";
RL   FEMS Microbiol. Lett. 75:225-230(1992).
RN   [2]
RP   SEQUENCE OF 1-64 FROM N.A.
RC   STRAIN=Hobbs FT10;
RX   MEDLINE=94297488; PubMed=7764998;
RA   East A.K., Collins M.D.;
RT   "Conserved structure of genes encoding components of botulinum
RT   neurotoxin complex M and the sequence of the gene coding for the
RT   nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL   Curr. Microbiol. 29:69-77(1994).
RN   [3]
RP   SEQUENCE OF 634-1002 FROM N.A.
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B.; Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
CC       BOND OF SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92906; AAA23263.1; -.
DR   EMBL; S73676; AAC60475.1; -.
DR   EMBL; X70820; CAA50151.1; -.
DR   EMBL; X70816; CAA50147.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    436       BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT   CHAIN       437   1274       BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT   METAL       227    227       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    228    228       BY SIMILARITY.
FT   METAL       231    231       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    445       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1274 AA;  146709 MW;  5B99756A7438B921 CRC64;

  Query Match             29.5%;  Score 316;  DB 1;  Length 1274;
  Best Local Similarity   36.2%;  Pred. No. 1e-17;
  Matches   79;  Conservative   51;  Mismatches   62;  Indels   26;  Gaps   11;

Qy       1 IKVNNWDLFFSPSEDNFT-NDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPE 59
Db     446 IRVNNSELFFVASESSYNENDINTPKEIDDTTNLNNNYRN-NLD---EVILDYNSQTIPQ 501

Qy      60 NISIENLSSDIIGQLELMPNIERFPNG----KKYELDKYTMFHYLRAQEFEHGKSRIALT 115
Db     502 -ISNRTLNT-LVQDNSYVPRYD--SNGTSEIEEYDVVDFNVFFYLHAQKVPEGETNISLT 557

Qy     116 NSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIAD 175
Db     558 SSIDTALLEESK-DIFFSSEFIDTINKPVNAALFIDWISKVIRDFTTEATQKSTVDKIAD 616

Qy     176 ITIIIPYIGPALNI------GNMLYKDDF----VGALI 203
Db     617 ISLIVPYVGLALNIIIEAEKGN--FEEAFELLGVGILL 652
```

```
RESULT 8
TETX_CLOTE
ID   TETX_CLOTE     STANDARD;      PRT;  1314 AA.
AC   P04958;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin).
OS   Clostridium tetani.
OG   Plasmid.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87053814; PubMed=3536478;
RA   Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA   Weller U., Hudel M., Habermann E., Niemann H.;
RT   "Tetanus toxin: primary structure, expression in E. coli, and
RT   homology with botulinum toxins.";
RL   EMBO J. 5:2495-2502(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CN3911;
RX   MEDLINE=87040747; PubMed=3774547;
RA   Fairweather N.F., Lyness V.A.;
RT   "The complete nucleotide sequence of tetanus toxin.";
RL   Nucleic Acids Res. 14:7809-7812(1986).
RN   [3]
RP   SEQUENCE OF 742-1314 FROM N.A.
RX   MEDLINE=86085672; PubMed=3510187;
RA   Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT   "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT   fragment C in Escherichia coli.";
RL   J. Bacteriol. 165:21-27(1986).
RN   [4]
RP   PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX   MEDLINE=90201034; PubMed=2108021;
RA   Krieglstein K., Henschen A., Weller U., Habermann E.;
RT   "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT   in tetanus toxin.";
RL   Eur. J. Biochem. 188:39-45(1990).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=92037649; PubMed=1935979;
RA   Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT   "Limited proteolysis of tetanus toxin. Relation to activity and
RT   identification of cleavage sites.";
RL   Eur. J. Biochem. 202:41-51(1991).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93010948; PubMed=1396558;
RA   Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA   Montecucco C.;
RT   "Tetanus toxin is a zinc protein and its inhibition of
RT   neurotransmitter release and protease activity depend on zinc.";
RL   EMBO J. 11:3577-3583(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX   MEDLINE=97475217; PubMed=9334741;
RA   Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA   Sax M.;
RT   "Structure of the receptor binding fragment HC of tetanus
RT   neurotoxin.";
RL   Nat. Struct. Biol. 4:788-792(1997).
CC   -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC       BOND OF SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
CC       SYNAPTOBREVIN.
CC   -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC       YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC       AND ARE NON-TOXIC AFTER SEPARATION.
CC   -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC       GANGLIOSIDE RECEPTORS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04436; CAA28033.1; -.
DR   EMBL; M12739; AAA23282.1; -.
DR   EMBL; X06214; CAA29564.1; -.
DR   PIR; A25689; BTCLTN.
DR   PDB; 1AF9; 29-APR-98.
DR   PDB; 1A8D; 14-OCT-98.
DR   MEROPS; M27.001; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW   3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    456       TETANUS TOXIN LIGHT CHAIN.
FT   CHAIN       457   1314       TETANUS TOXIN HEAVY CHAIN.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    233    233       BY SIMILARITY.
FT   METAL       236    236       ZINC (CATALYTIC) (BY SIMILARITY).
FT   TRANSMEM    226    246       POTENTIAL.
FT   TRANSMEM    669    689       POTENTIAL.
FT   DISULFID    438    466       INTERCHAIN.
FT   DISULFID   1076   1092
SQ   SEQUENCE   1314 AA;  150550 MW;  134C3657133EF81D CRC64;

  Query Match             29.1%;  Score 312;  DB 1;  Length 1314;
  Best Local Similarity   33.2%;  Pred. No. 2.2e-17;
  Matches   74;  Conservative   44;  Mismatches   67;  Indels   38;  Gaps    4;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||: | || |   :::|: :  : | :: :|  :    | ||| |   |
Db     467 IKIKNEDLTFIAEKNSFSEEPFQDEIVSYNTKNKPLNFNYSLDKIIVDY---------- 515

Qy      61 ISIENLSSDIIGQLELMPNIERFPNGKK----------------YELDKYTMFHYLRAQ 103
             || | |      :||    |  |                  : :|  |:: || ||
Db     516 ----NLQSKI-----TLPNDRTTPVTKGIPYAPEYKSNAASTIEIHNIDDNTIYQYLYAQ 566

Qy     104 EFEHGKSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDE 163
           :      || :|||||::||:| :::|::| |  : |||:  :  :|| ||  :: |||:|
Db     567 KSPTTLQRITMTNSVDDALINSTKIYSYFPS-VISKVNQGAQGILFLQWVRDIIDDFTNE 625

Qy     164 TSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           :|: :| |||:|:: |:|||||||||    |: :|:||| :|
Db     626 SSQKTTIDKISDVSTIVPYIGPALNIVKQGYEGNFIGALETTG 668
```

```
RESULT 9
BXD_CLOBO
ID   BXD_CLOBO      STANDARD;      PRT;  1276 AA.
AC   P19321;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
DE   (Bontoxilysin D).
GN   BOTD.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BVD/-3;
RX   MEDLINE=91016853; PubMed=2216736;
RA   Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
RA   Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
RT   "Nucleotide sequence of the gene encoding Clostridium botulinum
RT   neurotoxin type D.";
RL   Nucleic Acids Res. 18:5556-5556(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CB16;
RX   MEDLINE=93042276; PubMed=1420572;
RA   Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
RT   "The complete amino acid sequence of the Clostridium botulinum type D
RT   neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT   phage d-16 phi genome.";
RL   J. Vet. Med. Sci. 54:905-913(1992).
RN   [3]
RP   PARTIAL SEQUENCE.
RC   STRAIN=D-SA, and D-1873;
RX   MEDLINE=89339741; PubMed=2668193;
RA   Moriishi K., Syuto B., Kubo S., Oguma K.;
RT   "Molecular diversity of neurotoxins from Clostridium botulinum type D
RT   strains.";
RL   Infect. Immun. 57:2886-2891(1989).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-|-LEU-61 BOND OF
CC       SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
CC       STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC       BACTERIOPHAGE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54254; CAA38175.1; -.
DR   EMBL; S49407; AAB24244.1; -.
DR   PIR; S11455; S11455.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    442       BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
FT   CHAIN       443   1276       BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    437    450       INTERCHAIN (PROBABLE).
FT   VARIANT      15     16       ND -> PV (IN STRAIN D-SA).
FT   VARIANT      17     18       ND -> LQ (IN STRAIN D-1873).
FT   VARIANT     452    452       K -> Q (IN STRAIN D-SA).
FT   VARIANT     457    457       R -> T (IN STRAIN D-SA).
FT   VARIANT     457    457       R -> F (IN STRAIN D-1873).
FT   VARIANT     462    462       A -> D (IN STRAIN D-1873).
FT   VARIANT     489    489       K -> N (IN STRAIN CB16).
FT   VARIANT     644    644       N -> K (IN STRAIN CB16).
FT   VARIANT    1122   1122       Q -> R (IN STRAIN CB16).
SQ   SEQUENCE   1276 AA;  146871 MW;  C1EC50F46C8233E2 CRC64;

  Query Match             27.1%;  Score 290.5;  DB 1;  Length 1276;
  Best Local Similarity   32.6%;  Pred. No. 1.1e-15;
  Matches   71;  Conservative   39;  Mismatches   79;  Indels   29;  Gaps    5;

Qy       1 IKVNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPEN 60
           ||| |  | :   :|: : :: : : || :||::   :  |||   :  |       ||
Db     451 IKVKNNRLPYVADKDSISQEIFENKILTDETNVQNYSDKFSLD---ESILDGQVPINPEI 507

Qy      61 ISIENLSSDIIGQLELMPNIERFP-----------NGKKYELDKYTMFHYLRAQEFEHG 108
           :                |:||:   |           :  || :|    ::|| :|:  :
Db     508 VD-----------PLLPNVNMEPLNLPGEEIVFYDDITKY-VDYLNSYYYLESQKLSNN 554

Qy     109 KSRIALTNSVNEALLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVS 168
             | || || |||   :::|||  |   :|||  :| :|| |  ::| |||    :
Db     555 VENITLTTSVEEALGYSNKIYTFLPS-LAEKVNKGVQAGLFLNWANEVVEDFTTNIMKKD 613

Qy     169 TTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSG 206
           | |||:|:::||||||||||||||   : :|  |   :|
Db     614 TLDKISDVSVIIPYIGPALNIGNSALRGNFNQAFATAG 651


RESULT 10
BXC1_CLOBO
ID   BXC1_CLOBO     STANDARD;      PRT;  1290 AA.
AC   P18640;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE   (Bontoxilysin C1).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90370487; PubMed=2204031;
RA   Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
```

```
RA   Boquet P., Popoff M.R.;
RT   "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL   Nucleic Acids Res. 18:4924-4924(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=91024998; PubMed=2222445;
RA   Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA   Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT   "The complete nucleotide sequence of the gene coding for botulinum
RT   type C1 toxin in the C-ST phage genome.";
RL   Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN   [3]
RP   SEQUENCE OF 2-25.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=88153072; PubMed=2450068;
RA   Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
RA   Oguma K.;
RT   "Establishment of a monoclonal antibody recognizing an antigenic site
RT   common to Clostridium botulinum type B, C1, D, and E toxins and
RT   tetanus toxin.";
RL   Infect. Immun. 56:898-902(1988).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94038966; PubMed=7901002;
RA   Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT   "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
RT   cleaving HPC-1/syntaxin.";
RL   EMBO J. 12:4821-4828(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC       HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC       WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC       FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC       BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC   -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
CC       STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC       BACTERIOPHAGE.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66433; CAA47060.1; -.
DR   EMBL; X72793; CAA51313.1; -.
DR   EMBL; X53751; CAA37780.1; -.
DR   EMBL; D90210; BAA14235.1; -.
DR   EMBL; X62389; CAA44263.1; -.
DR   PIR; S11291; S11291.
DR   PIR; A35396; A35396.
DR   PIR; A43503; A43503.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    448       BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
FT   CHAIN       449   1290       BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
FT   METAL       228    228       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    229    229       BY SIMILARITY.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    452       INTERCHAIN (PROBABLE).
FT   CONFLICT     84     84       P -> T (IN REF. 2).
SQ   SEQUENCE   1290 AA;  148734 MW;  71FBE379F97129E8 CRC64;

  Query Match             24.6%;  Score 263;  DB 1;  Length 1290;
  Best Local Similarity   31.2%;  Pred. No. 1.9e-13;
  Matches   64;  Conservative   39;  Mismatches   94;  Indels    8;  Gaps    4;

Qy       3 VNNWDLFFSPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPENIS 62
           | | || |     :   |:   ::|  :| :    :|:|:|   |  |:    |  |:
Db     457 VKNTDLPFIGDISDVKTDIFLRKDINEETEVIYYPDNVSVD---QVILS---KNTSEHGQ 510

Qy      63 IENLSSDIIGQLELMPNIER-FPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEA 121
           :: |   |  : |::|   : | : :   :|    ::|| :|:          | |: ||
Db     511 LDLLYPSIDSESEILPGENQVFYDNRTQNVDYLNSYYYLESQKLSDNVEDFTFTRSIEEA 570

Qy     122 LLNPSRVYTFFSSDYVKKVNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIP 181
           | | ::|||:| :    |||   :  :|| |   :| |||       | |||:|:: |||
Db     571 LDNSAKVYTYFPT-LANKVNAGVQGGLFLMWANDVVEDFTTNILRKDTLDKISDVSAIIP 629

Qy     182 YIGPALNIGNMLYKDDFVGALIFSG 206
           |||||||| | : : :|  |   :|
Db     630 YIGPALNISNSVRRGNFTEAFAVTG 654


RESULT 11
Y402_BUCAI
ID   Y402_BUCAI     STANDARD;      PRT;   246 AA.
AC   P57482;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein BU402.
GN   BU402.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tokyo 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- SIMILARITY: BELONGS TO THE UPF0169 (COML) FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001119; BAB13105.1; -.
DR   InterPro; IPR005156; UPF0169.
DR   Pfam; PF03696; UPF0169; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   246 AA;  29937 MW;  1289ED74CD94088F CRC64;

  Query Match             10.0%;  Score 107;  DB 1;  Length 246;
  Best Local Similarity   21.8%;  Pred. No. 0.088;
  Matches   44;  Conservative   40;  Mismatches   74;  Indels   44;  Gaps    8;
```

```
Qy      14 EDNFTNDLNKGEEITSDTN-IEAAEENISLDLIQQYYLTFNFDNEPENIS---------- 62
           ::|| | ::  |:|  : |    : : | :|||  ||   |||   :||
Db      46 KENFDNAISILEKIKKNNNTANISNDKIQIDLIYAYYKILNFDQARKNIEEFMYFYPNHP 105

Qy      63 ----------IENLSSDIIGQLELMPNIERFPNGKKYELD-----KYTMFHYLRAQEFEH 107
                   : ::| |      : | |  : |   |  :     || :: | :::   :
Db     106 NIDYVVYIQCLISMSLDKNRFFSVFP-INYYKNDYFYAKNAFFQLKYFIYQYPKSRYVVN 164

Qy     108 GKSR-IALTNSVNEALLNPSRVYTFFSSDYVKKVNKATE----------AAMFLGWVEQ- 155
           |   | : | ::|  |:  : | ||  :|:  :|:  |          |   | ::|:
Db     165 AKKNLIYIKNRLSEHDLSILKFY-FFHKEYIAVINRGEEMLQRYSETPSARKALIYIEKS 223

Qy     156 ----LVYDFTDETSEVSTTDKI 173
               ::|   : |::   :||
Db     224 YYALKIFDTAKKISKIILLNKI 245


RESULT 12
BXCN_CLOBO
ID   BXCN_CLOBO     STANDARD;      PRT;  1196 AA.
AC   P46081;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Botulinum neurotoxin type C1, nontoxic component.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=92231894; PubMed=1567404;
RA   Tsuzuki K., Kimura K., Fujii N., Yokosawa N., Oguma K.;
RT   "The complete nucleotide sequence of the gene coding for the
RT   nontoxic-nonhemagglutinin component of Clostridium botulinum type C
RT   progenitor toxin.";
RL   Biochem. Biophys. Res. Commun. 183:1273-1279(1992).
CC   -!- FUNCTION: THE NONTOXIC COMPONENT IS NECESSARY TO MAINTAIN
CC       TOXICITY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62389; CAA44262.1; -.
DR   InterPro; IPR000395; Bontoxilysin.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
KW   Neurotoxin.
SQ   SEQUENCE   1196 AA;  138740 MW;  4BD5956274D7F9C3 CRC64;

  Query Match             9.9%;  Score 106;  DB 1;  Length 1196;
  Best Local Similarity   20.4%;  Pred. No. 0.77;
  Matches   50;  Conservative   48;  Mismatches   81;  Indels   66;  Gaps   11;

Qy      12 PSEDNFTNDLNKGEEITSDTNIEAAEENISL--------------DLIQQYYLTFNFD- 55
           |  :  || ::| |:: :  |    | ||||             |    | : :| :
Db     387 PLSNKNTNIISKPEKVVNLVN----ENNISLMKSNIYGDGLKGTTEDFYSTYKIPYNEEY 442

Qy      56 ----NEPENISIENLSSDIIGQLELMPNIERFPNGK----------KYELDKYT--MFH 98
               |: :|  : |:| : :  :  : :|  : :              |:  :|    :
Db     443 EYRFNDSDNFPLNNISIEEVDSIPEIIDINPYKDNSDNLVFTQITSMTEEVTTHTALSIN 502

Qy      99 YLRAQEFEHGKSRIALTNSVNEAL----------LNPSRVYTFFSS--DYVKKVNK---- 142
           ||:||          :||: |  |          : | ||:|  :   |:: :
Db     503 YLQAQ----------ITNNENFTLSSDFSKVVSSKDKSLVYSFLDNLMSYLETIKNDGPI 552

Qy     143 ATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGAL 202
             |:  :| |:::: ::: : :     | :  |  :: : | |||| |    : ||
Db     553 DTDKKYYL-WLKEVFKNYSFDINLTQEIDSMCGINEVVLWFGKALNILNT--SNSFVEEY 609

Qy     203 IFSGA 207
           |||
Db     610 QDSGA 614


RESULT 13
VA0D_YEAST
ID   VA0D_YEAST     STANDARD;      PRT;   345 AA.
AC   P32366;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Vacuolar ATP synthase subunit d (EC 3.6.3.14) (V-ATPase d subunit)
DE   (Vacuolar proton pump d subunit) (V-ATPase 39 kDa subunit) (V-ATPase
DE   subunit M39).
GN   VMA6 OR YLR447C OR L9324.8.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=93286119; PubMed=8509410;
RA   Bauerle C., Ho M.N., Lindorfer M.A., Stevens T.H.;
RT   "The Saccharomyces cerevisiae VMA6 gene encodes the 36-kDa subunit of
RT   the vacuolar H(+)-ATPase membrane sector.";
RL   J. Biol. Chem. 268:12749-12757(1993).
CC   -!- FUNCTION: VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY
CC       OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS. THE ACTIVE
CC       ENZYME CONSISTS OF A CATALYTIC V1 DOMAIN ATTACHED TO AN INTEGRAL
CC       MEMBRANE V0 PROTON PORE COMPLEX. THIS SUBUNIT IS A NON-INTEGRAL
CC       MEMBRANE COMPONENT OF THE MEMBRANE PORE DOMAIN AND IS REQUIRED FOR
CC       PROPER ASSEMBLY OF THE V0 SECTOR. MIGHT BE INVOLVED IN THE
CC       REGULATED ASSEMBLY OF V1 SUBUNITS ONTO THE MEMBRANE SECTOR OR
CC       ALTERNATIVELY MAY PREVENT THE PASSAGE OF PROTONS THROUGH V0 PORES.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
CC       peripheral catalytic V1 complex (components A to H) attached to
CC       an integral membrane V0 proton pore complex (components: a, c, c',
CC       c'' and d).
CC   -!- SIMILARITY: BELONGS TO THE V-ATPASE V0D/AC39 SUBUNIT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L11584; AAA35210.1; -.
DR   EMBL; U22382; AAB67533.1; -.
DR   PIR; S35105; S35105.
DR   PIR; A45994; A45994.
DR   SGD; S0004439; VMA6.
DR   InterPro; IPR002843; ATPsynt_AC39sub.
DR   Pfam; PF01992; vATP-synt_AC39; 1.
KW   Hydrolase; Hydrogen ion transport.
FT   CONFLICT     32     32       N -> T (IN REF. 1).
SQ   SEQUENCE   345 AA;  39790 MW;  53A19450CAF35632 CRC64;

  Query Match             9.2%;  Score 98;  DB 1;  Length 345;
  Best Local Similarity   29.0%;  Pred. No. 0.71;
  Matches   36;  Conservative   17;  Mismatches   41;  Indels   30;  Gaps    6;

Qy      11 SPSEDNFTNDLNKGEEITSDTNIEAAEENISLDLIQQYY-----------------LTF 52
           :|    | | :  ||:  | |||     :    :: :|                  | |
```

```
Db    156 TPLAPYFKNCFDTAEEL-DDMNIEIIRNKLYKAYLEDFYNFVTEEIPEPAKECMQTLLGF 214

Qy     53 NFDNEPENISIENL-SSDIIGQL--ELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHGK 109
          |    ||:: :| |||   |  :|:|||     || | |     ||  :||:||  :
Db    215 EADRRSINIALNSLQSSDIDPDLKSDLLPNI-----GKLYPL---ATFHLAQAQDFEGVR 266

Qy    110 SRIA 113
          : :|
Db    267 AALA 270


RESULT 14
MT18_YEAST
ID   MT18_YEAST     STANDARD;      PRT;  1032 AA.
AC   P40469; P89106;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   DNA repair/transcription protein MET18/MMS19.
GN   MET18 OR MMS19 OR YIL128W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RX   MEDLINE=97098656; PubMed=8943333;
RA   Lauder S., Bankmann M., Guzder S.N., Sung P., Prakash L.,
RA   Prakash S.;
RT   "Dual requirement for the yeast MMS19 gene in DNA repair and RNA
RT   polymerase II transcription.";
RL   Mol. Cell. Biol. 16:6783-6793(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RA   Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA   Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA   Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA   Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA   Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA   Walsh S.V., Whitehead S.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-161 FROM N.A.
RC   STRAIN=S288c / AB972;
RA   Churcher C., Barrell B.G., Rajandream M.A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 162-1032 FROM N.A.
RC   STRAIN=S288c / AB972;
RA   Hamlyn N., Churcher C., Barrell B.G., Rajandream M.A.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN NUCLEOTIDE EXCISION REPAIR (NER) AND RNA
CC       POLYMERASE II (POL II) TRANSCRIPTION. IT PROBABLY DOES NOT
CC       PARTICIPATE DIRECTLY IN NER AND POL II TRANSCRIPTION BUT EXERTS
CC       ITS BIOLOGICAL EFFECTS BY INFLUENCING THE ACTIVITY OF TFIIH AND
CC       POSSIBLY OTHER DNA REPAIR AND TRANSCRIPTION FACTORS AS AN UPSTREAM
CC       REGULATORY ELEMENT. INVOLVED IN SPORULATION AND POSSIBLY IN THE
CC       RAD52 RECOMBINATIONAL REPAIR PATHWAY.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DOMAIN: HAS 15 TANDEM LEUCINE RICH REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U70559; AAB38865.1; -.
DR   EMBL; Z38059; CAA86150.1; -.
DR   EMBL; Z46833; CAA86864.1; -.
DR   TRANSFAC; T03493; -.
DR   SGD; S0001390; MET18.
KW   DNA repair; Nuclear protein; Repeat.
FT   CONFLICT    230    230       P -> Q (IN REF. 1).
FT   CONFLICT    329    329       D -> G (IN REF. 1).
FT   CONFLICT    335    335       V -> M (IN REF. 1).
FT   CONFLICT    361    361       V -> I (IN REF. 1).
SQ   SEQUENCE   1032 AA;  117882 MW;  D05DCC488E098814 CRC64;

  Query Match             8.9%;  Score 95.5;  DB 1;  Length 1032;
  Best Local Similarity   22.1%;  Pred. No. 4.5;
  Matches   46;  Conservative   33;  Mismatches   76;  Indels   53;  Gaps   10;

Qy      3 VNNWDLFFSP----------SEDNFTNDLNK---GEEITSDTNIEAAEENISLDLIQQY 48
          : || | ::            :|:   |  ||    :::   || :|  : |:|  :| |
Db    297 LENWTLLWNALKFEIMQNSEGNENTLLNPYNKDQQSDDVGQYTNYDACLKIINLMALQLY 356

Qy     49 YLTFNFDNEPENISIENLSSDIIGQLELMPNIERFPNGKKYELDKYTMFHYLRAQEFEHG 108
            |||     :| |   : ::   || ||        ||| |       | |     |
Db    357 ----NFD----KVSFEKFFTHVLD--ELKPNF-------KYEKDLKQTCQILSA----IG 395

Qy    109 KSRIALTNSVNEA-----LLNPSRV---------YTFFSSDYVKKVNKATEAAMFLGWVE 154
          : : | |  :      |:| | |           ::||   |:    : :: ::
Db    396 SGNVEIFNKVISSTFPLFLINTSEVAKLKLLIMNFSFFVDSYIDLFGRTSKESLGTPVPN 455

Qy    155 QLVYDFTDE----TSEVSTTDKIADITI 178
          : :: ||     |   |    |::||
Db    456 NKMAEYKDEIIMILSMALTRSSKAEVTI 483


RESULT 15
Y198_RICPR
ID   Y198_RICPR     STANDARD;      PRT;   355 AA.
AC   Q9ZDW7;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein RP198 precursor.
GN   RP198.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madrid E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ235270; CAA14663.1; -.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    355       HYPOTHETICAL PROTEIN RP198.
SQ   SEQUENCE   355 AA;  40780 MW;  A08B7655BECD86E0 CRC64;

  Query Match             8.7%;  Score 93.5;  DB 1;  Length 355;
  Best Local Similarity   26.4%;  Pred. No. 1.7;
  Matches   57;  Conservative   32;  Mismatches   76;  Indels   51;  Gaps   14;

Qy      1 IKVNNWDL-FFSPSEDNF-TNDLNKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEP 58
```

```
          ||:||  |  |  ::: |  |||     :    |: |::||||| |   ||      :|
Db    146 IKLNNKTLSIFKDNQEIFKINDL---AFLIKKHNL-ASQENISLFLNMHYY------SEK 195

Qy     59 ENISIENLSSDIIGQLELMP----------NIERF----PNGKKYELDKYTMFHYLRAQ 103
          : :: :| : |:    :              |||||     |  |  |:    |    |
Db    196 DILNFKNANLDMATSFKFAKNGKDAAILENLNIERFIFTCDNESKVNLNGTLQFF---AN 252

Qy    104 EFEHGKSRIALT--NSVNEALLNPSRVYTFFSSDYVKKV-----NKATEAAMFLGWVEQL 156
          :   |     |   ||: : || |: :   ||   :| :     || ::        |||
Db    253 KLPKGILSFELENYNSIVDKLL-PNSI--LFSKKTIKTIIAKAMNKTSD--------EQL 301

Qy    157 VYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNM 192
            ||:   |  : | :    | :    :|||:|
Db    302 ---NTDKNDTNSVYNNIKNAKFDIAFSDKGINIGSM 334
```

Search completed: March 13, 2003, 11:40:36
Job time : 13.5447 secs